A Research Report Published: **Nov 2023**

EPIC
Education Policy Innovation Collaborative

Michigan's 2022-23 Benchmark Assessments

ACKNOWLEDGEMENTS

The authors wish to acknowledge the many people who graciously gave of their time in support of this effort. We are especially grateful to our partners for their collaboration and thoughtful feedback. In particular, we would like to thank Dr. Sue Carnell, Kate Cermak, Dr. Delsa Chapman, David Judd, Ann-Marie Mapes, Andy Middlestead, Gerry Polverento, and Dr. Michael Rice from the Michigan Department of Education; Tom Howell and Lauren Paluta from the Center for Educational Performance and Information; Kyle Kwaiser and Nicole Wagner from the Michigan Education Data Center at the University of Michigan; and Don Dailey from the Michigan Data Hub. At Michigan State University, we thank Emily Mohr and Meg Turner for coordinating and facilitating the project, Michelle Huhn for her assistance with formatting, and Scott Imberman for his helpful feedback. Finally, we thank Bryant Hopkins and Katharine Strunk for their contributions in earlier stages of this work and Bridgette Redman for her excellent copy-editing.

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.

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.

Table of Contents

Executive Summary

BACKGROUND AND PURPOSE

In order to monitor student progress toward learning goals in the wake of the COVID-19 pandemic, the Michigan legislature mandated benchmark assessment testing for all K-8 students in both the fall and spring of each school year starting in 2020-21 (2020 PA 149, 2021 PA 48). To help interpret and contextualize the assessment results, the Education Policy Innovation Collaborative (EPIC) has partnered with the Michigan Department of Education (MDE), the Center for Educational Performance and Information (CEPI), the Michigan Data Hub (MDH), and the Michigan Education Data Center (MEDC) to prepare and deliver a series of reports to the governor and the Senate and House standing committees responsible for education legislation in the Michigan legislature. This is the fifth report in this series.

In past reports, we showed that Michigan students experienced more math and reading growth in 2021-22 than in 2020-21, but that these improvements were not enough to counteract the effects of unfinished learning in 2020-21. Students with less access to in-person instruction experienced less achievement growth than those with more access, though these gaps began to diminish for students who returned to in-person learning in 2021-22. Sociodemographic achievement gaps worsened over the course of the 2020-21 school year and remained stable in 2021-22.

RESEARCH QUESTIONS AND STUDY OVERVIEW

In this report, we use newly available data from the 2022-23 school year to expand on our previous analyses and further investigate each of the following questions:

1. **How do Michigan students' achievement trajectories in recent years compare to pre-pandemic trends?** To better understand how the COVID-19 pandemic has affected Michigan students and the extent to which they have recovered academically, we examine changes in Michigan students' benchmark assessment scores across the fall and spring semesters of each school year and compare these trends to national and state-specific norms for the same assessments from pre-pandemic.
2. **How did Michigan students' growth over the course of each year compare with typical yearly growth before the COVID-19 pandemic?** We compare students' growth between their fall and spring benchmark assessments each

year to pre-pandemic national norms for each assessment, subject, and grade level, which differ based on students' baseline achievement in the fall.

3. **How have trends in achievement and growth differed across subgroups of Michigan students?** We compare patterns in student achievement and growth across sociodemographic subgroups and instructional modalities (i.e., in-person, hybrid, or remote).

Our analyses include benchmark assessment results from approximately 773,000 of Michigan's 947,000 K-8 students in 769 of Michigan's 852 school districts. While these analyses help to deepen our understanding of Michigan public school students' academic achievement and growth between fall 2020 and spring 2023, they are based on imperfect and incomplete data. For instance, the students who were affected the most by the COVID-19 pandemic may have also been less likely to participate in benchmark assessments and more likely to switch districts over the course of our study period and may therefore be underrepresented in our analyses.

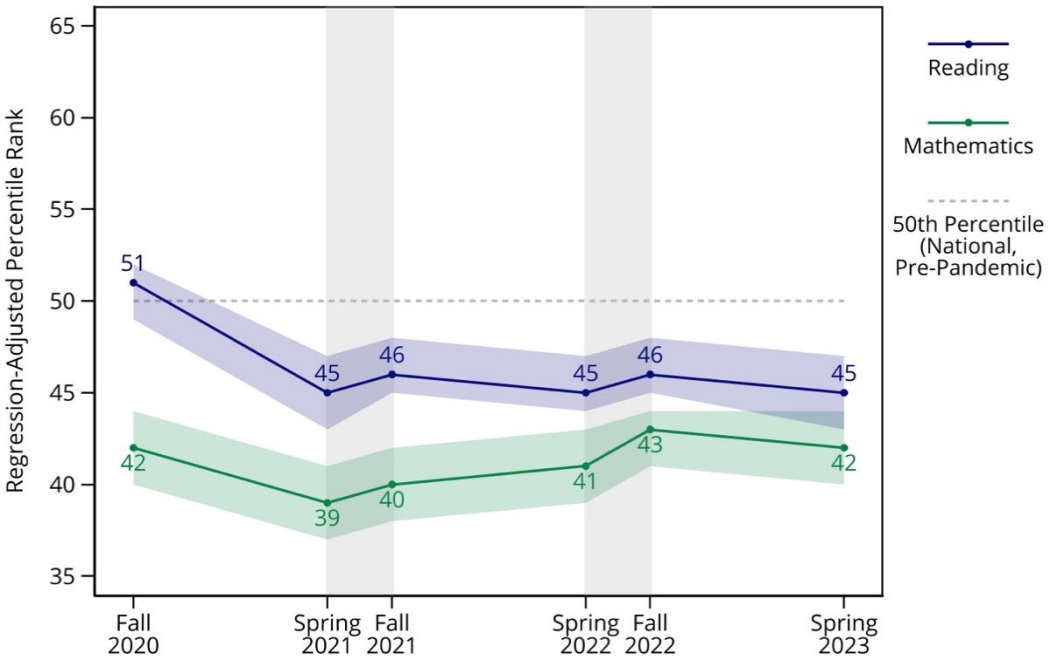
Michigan's benchmark assessment legislation allows districts to choose an appropriate assessment from one of four MDE-approved providers, and thousands of students participated in assessments from each of the four. More than 600 districts participated in an NWEA MAP Growth assessment in 2022-23, while 75 districts administered Curriculum Associates' i-Ready Diagnostic assessments, 83 districts participated in Renaissance Learning's Star 360 assessments, and 24 administered one of DRC's benchmark assessments. These assessments are all designed in slightly different ways, cover slightly different content, and tend to appeal to different types of districts. We analyze data from each assessment separately to identify common themes in their results and ways that they differ. Where possible, we also use metrics that have similar meanings across assessments to summarize Michigan students' benchmark assessment performance overall.

KEY FINDINGS

On Average, Math Achievement Has Improved Slightly Since Spring 2021, While Reading Achievement Has Stayed About the Same

After accounting for differences across assessments, grade levels, and the demographic composition of students in each district, we find evidence of slight improvements in Michigan students' math achievement but very little change in reading achievement since 2021-22. As Figure I shows, average scores for Michigan students in fall 2020 were at about the 42nd percentile of the national norming distribution for each assessment, declining to the 39th percentile by spring 2021, and eventually returning to the 42nd percentile by spring 2023.

Figure I. Regression-Adjusted Percentile Ranks by Semester (MAP Growth & i-Ready, 3rd-8th Grade)



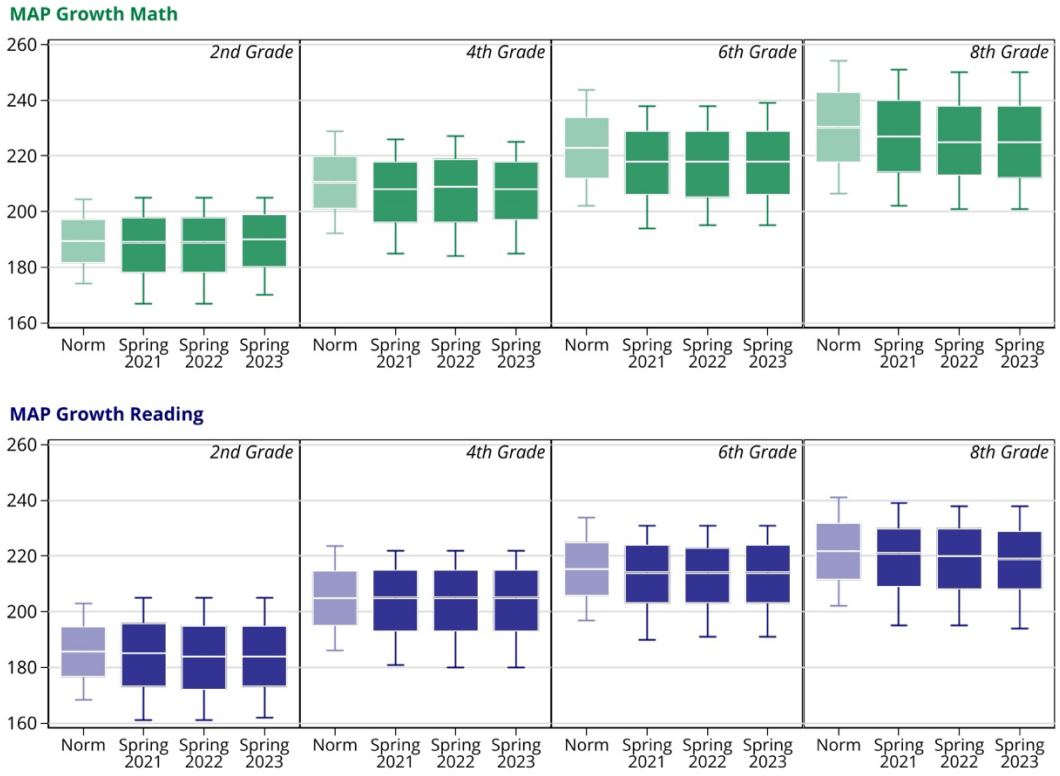
Notes: These regression estimates are based on district grade-average scores across students with MAP Growth or i-Ready scores in every possible testing period. We standardized scores relative to pre-pandemic national norms for each assessment and converted all estimates into percentile ranks. Models control for student demographics.

This suggests that Michigan students' math achievement, relative to other students across the country pre-pandemic, has recovered to about the same level as it was in fall 2020. However, since scores probably fell at the beginning of the COVID-19 pandemic, this is still likely below where scores were before the initial school closures in 2019-20. Reading scores also declined during the 2020-21 school year, falling from the 51st to the 45th percentile of the national distribution between fall 2020 and spring 2021, and have not changed substantially since then.

Michigan Students' Achievement Levels Vary to a Greater Extent Than Would Have Been Expected Pre-Pandemic

Gaps between Michigan's highest and lowest scoring students are larger than those of students across the country who took the same assessments pre-pandemic. Figure II shows the distribution of MAP Growth math and reading scores, which generally reflect the same patterns we observed for other grades and vendors.

Figure II. Distribution of Spring Scale Scores by School Year and Grade Level, MAP Growth



Notes: Each vertical line shows the 10th and 90th percentiles of students' spring scores, while each rectangle shows the 25th, 50th, and 75th percentiles. "Norm" represents the pre-pandemic national norm.

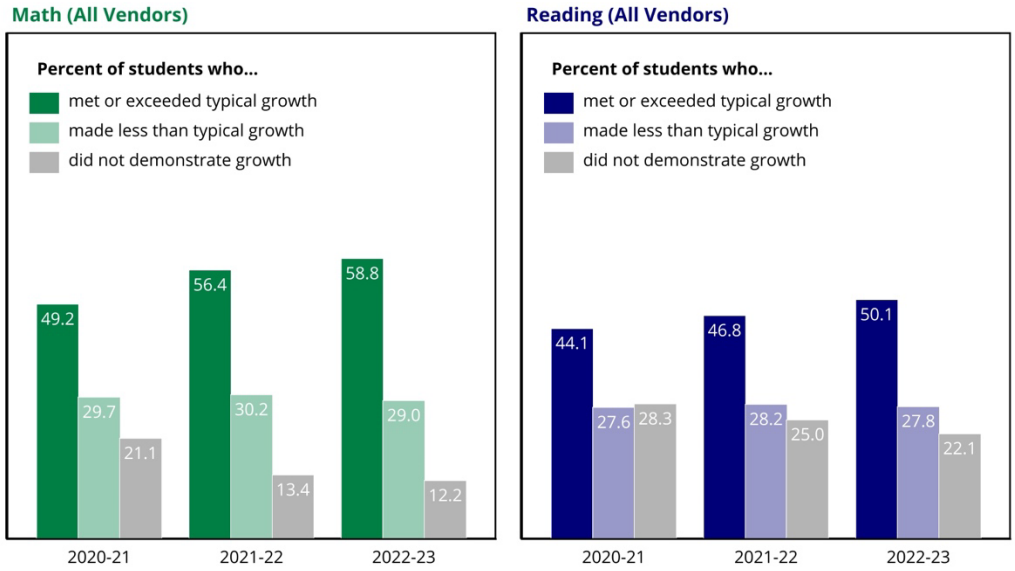
We find that, in nearly all cases, the distribution of Michigan students' scores in each pandemic-affected testing period is wider than the national pre-pandemic distribution, suggesting that there is more variation in achievement across Michigan students than in the nationally representative norming samples. Another consistent pattern is that the 90th percentile of Michigan students is only slightly below the 90th percentile of the national norming distribution, whereas the 10th percentile of Michigan students is very far below the 10th percentile of the national distribution. This suggests that Michigan's lowest-scoring students were disproportionately affected by disrupted learning during the COVID-19 pandemic. These gaps between Michigan's higher and lower-scoring students remained consistent across all six testing periods.

Students Were More Likely to Reach Targets for "Typical" Growth in 2022-23, but Many Continued to Struggle

We define "typical" yearly growth as the median increase in scale scores for students who took the same tests before the COVID-19 pandemic and had similar baseline scores. This means that we would expect about 50% of students to reach or exceed

these targets in a typical year. We would generally expect to see very few students who do not demonstrate any growth at all (i.e., whose scale scores decrease or do not change) between their fall and spring benchmark assessments.

Figure III. Fall-to-Spring Growth Outcomes by School Year (All Vendors)



Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" represent the median growth for students before the pandemic who took the same benchmark assessments in the subject area and grade level and had similar initial scores in the fall.

In 2020-21, Michigan students were less likely to achieve a "typical" year's growth and more likely to demonstrate no growth at all, compared to similar students pre-pandemic. As Figure III shows, these rates improved in each subsequent school year. In 2022-23, more than 50% of students reached or exceeded targets for "typical" yearly growth, suggesting that, on average, student learning returned to or surpassed pre-pandemic rates. However, for students who were already behind at the beginning of the year, a "typical" year's learning gains would not be enough to "catch up." As we saw in Figure I, learning rates have not accelerated enough for Michigan students' average scores to reach pre-pandemic norms. There are also still substantial percentages of students—12% in math and 22% in reading—who did not demonstrate growth at all in 2022-23.

Groups of Districts and Students That Were Most Negatively Affected by the COVID-19 Pandemic Also Experienced the Most Learning Recovery, But Some Remain Behind

Districts that operated in a remote or hybrid format for part or all of 2020-21 primarily drove the overall improvements in student growth outcomes since then. These districts are more likely to be in urban areas, serve more diverse student populations, and have more students from economically disadvantaged backgrounds. Accelerated rates of learning in these districts led to improvements in overall achievement and growth outcomes at the state level, and to improvements in achievement gaps. As we showed in previous reports, some achievement gaps across races/ethnicities and students' economically disadvantaged status worsened throughout the 2020-21 school year. Following disproportionate improvements among these same subgroups of students in subsequent years, most of these achievement gaps have improved since spring 2021 and some have returned to their fall 2020 levels.

SUMMARY

This report shows that, following more than three years of unprecedented disruptions to learning and schooling, student achievement trends in Michigan show signs of progress in some areas but still have a long way to go. On average, since spring 2021, student achievement has improved slightly in math but changed very little in reading. While the majority of students achieved a typical year's growth on their benchmark assessments in 2022-23, they would need to learn at a more accelerated rate to catch up to pre-pandemic achievement levels. Moreover, gaps between Michigan's highest- and lowest-performing students are larger than would have been expected before the COVID-19 pandemic, and some students are falling further behind rather than catching up. On the other hand, the same subgroups of students and districts that experienced the greatest declines in early stages of the pandemic also experienced the most recovery, and some of the achievement gaps across demographic groups and instructional modalities that widened in 2020-21 have since improved.

We must place all these results in the context of the imperfect data available to analyze student learning growth during the COVID-19 pandemic. Not only were participation rates lower than a typical end-of-year summative assessment, but the resulting analytic samples are not entirely reflective of Michigan's larger student population. In addition, we document a likely "at-home advantage" for early elementary students in fall 2020, which makes assessing growth and adherence to expected growth trajectories in the 2020-21 school year difficult for young cohorts of students. Nonetheless, the results presented herein provide important information for policymakers, educators, and stakeholders as we continue to grapple with the academic effects of the pandemic on Michigan's students.

Section One: Introduction

As is well established by now, the COVID-19 pandemic has severely affected student achievement across the United States and the effect has been particularly acute for certain student subgroups. This includes students of color, those receiving additional services, students attending high poverty schools, students attending elementary schools, those who learned remotely, and those with lower baseline achievement (Bailey et al., 2021; Dorn et al., 2020; Goldhaber, Kane, McEachin, & Morton, 2022; Goldhaber, Kane, McEachin, Morton, et al., 2022; Pier et al., 2021). Summative assessment data from more than 20 states consistently show marginal improvements in math in 2022-23, and varied rates of recovery (or in some cases, further declines) in ELA (Barnum & Belsha, 2023; *State Test Score Data Briefs*, 2023).

Previously, to establish educational goals and monitor student progress under such unprecedented circumstances, Michigan Governor Gretchen Whitmer signed the "Return to Learn" package of bills into law in 2020 (2020 PA 147, 148, 149; for a more in-depth discussion of these laws, see Michigan's 2020-21 and 2021-22 Benchmark Assessments Report). Therefore, districts were required to select and administer appropriate benchmark assessments—one of four state-approved assessments or a locally developed alternative that met specific criteria—to all K-8 students at the beginning and end of the school year to track student achievement during this time and then report these results through the Michigan Data Hub (MDH) network for use in a statewide aggregate report. In 2021, this requirement was extended to include additional school years (2021 PA 48).

CONTEXT AND PURPOSE OF THIS REPORT

This report is the fifth in a series that the Education Policy Innovation Collaborative (EPIC) at Michigan State University has provided to the Michigan Department of Education (MDE), the governor, and the House and Senate standing committees responsible for education legislation to give insight into Michigan students' progress toward learning goals in the wake of the COVID-19 pandemic. EPIC prepared this report in collaboration with MDE, the Center for Educational Performance and Information (CEPI), the MDH, and the Michigan Education Data Center (MEDC) at the University of Michigan as a summary of the student academic growth across the 2020-21, 2021-22, and 2022-23 school years.

The first four reports in this series, released in August 2021, January 2022, April 2022, and October 2022, examined student progress toward learning goals at various times throughout the first two full school years of the COVID-19 pandemic—2020-21 and 2021-22. Overall, these reports found that Michigan students experienced more math and reading growth in 2021-22, though often not enough to counteract the effects of unfinished learning in 2020-21. Among the students who made less than typical growth, average learning gains were greater in 2021-22 than the prior year. Further, while historically underserved populations of students remained less likely to achieve a typical year's growth, disparities between subgroups decreased over time. Finally, students with less access to in-person instruction in 2020-21 experienced less achievement growth, but gaps shrunk substantially once most districts returned to in-person learning in 2021-22.

This fifth report extends our analyses through the end of the 2022-23 school year. Specifically, in this analysis, we examine achievement trajectories and growth over the past six semesters—from fall 2020 to spring 2023—and assess differences in performance across subgroups of students with different demographic characteristics and those who participated in different modes of instruction (e.g., fully in-person, fully remote, or hybrid instruction) in 2020-21.

Section Two: Data and Methods

Each year, millions of K-12 students across the country participate in benchmark assessments. Benchmark assessments are designed to help educators and administrators track students' progress toward grade-level standards and learning goals, and to provide feedback to help drive future instruction.

Michigan's benchmark assessment legislation requires districts to administer to all K-8 students in the fall and spring of each school year a benchmark assessment either from the MDE-approved provider list, an assessment that provides adequate progress monitoring, or a local benchmark assessment. This requirement first went into effect in the 2020-21 school year and will continue through at least the end of 2023-24. Districts that chose an assessment from one of the four approved providers are required to provide results of these assessments through MDH.

In this section, we describe the analytic samples and methods we use in this report to understand student achievement in Michigan over the past three school years. For a full description of the unique characteristics of each MDE-approved benchmark assessment, please see the first report in this series.

PARTICIPATION AND ANALYSIS SAMPLES

Below, we provide details about the benchmark assessment data that districts submitted to the MDH and the sample of students we use to describe achievement in Michigan. For a full discussion of our general data exclusions, the sociodemographic and modality data that is merged with the benchmark outcomes, and our aggregate data file construction processes, please see our fourth report in this series.

District Participation

Under Michigan's benchmark assessment legislation, school districts serving K-8 students are expected to submit benchmark assessment data in some form. For this analysis, CEPI identified districts of interest as those that served students in at least one K-8 grade level and were open as of the official fall student count date for the 2022-23 school year (October 5, 2022) and remained open as of the official spring student count date (February 8, 2023).

In total, 772 of Michigan's 852 school districts provided some form of benchmark assessment data for the 2022-23 school year through the MDH. Of these, 750 provided student-level data, 22 provided aggregate files that they prepared themselves. We omitted 3 of these districts from our analyses because all the assessment results they provided were from time periods, grade levels, or subject areas that are not within the scope of this report. The remaining 769 districts are represented in at least some of the analyses in our report. This includes 643 districts using NWEA's MAP Growth, 75 using Curriculum Associates' i-Ready assessments, 83 using Renaissance Learning's Star 360 assessments, and 24 using DRC's ICAs or MDE's K-2s. Fifty-four of these districts administered assessments from two different providers and one district used assessments from three providers. These 769 districts teach 893,769 K-8 students, or 94% of the population of K-8 students in Michigan.

The Return to Learn legislation specifies a few options for districts as alternatives to the four approved benchmark assessment providers. The remaining 83 districts that did not provide any data through the MDH indicated through the Grant Electronic Monitoring System/Michigan Administrative Review System (GEMS/MARS; 2020 PA 149) that they either selected an alternate vendor or locally developed assessment, did not plan to submit any benchmark assessment data, did not provide the necessary authorization for MEDC and EPIC to access their data in the MDH, or provided the authorization but did not have any student benchmark assessment data in the MDH by the deadline for us to include them in the report.

Analysis Samples

Before aggregating the student-level benchmark assessment data that districts provided through the MDH, we restricted the sample to exclude: 1) districts that were not required to report data under Michigan's benchmark assessment legislation (i.e., districts that did not use products from an MDE-approved assessment provider and districts that opened after the official fall 2021 count date or closed before the official spring 2022 student count date); 2) students who were not in grades K-8; 3) results from assessments in subject areas other than math and reading/ELA or that only covered a narrow sub-topic within math or reading/ELA (i.e., the Smarter Balanced Interim Assessment Blocks, or IABs, which do not cover as broad a range of topics as the ICAs); and 4) results from assessments that were not normed for the grade level of the assessed student (i.e., results from Star Early Literacy assessments for students above grade 3 and results from Star Math assessments for students in kindergarten).

While the **full analytic sample** includes data from all students with valid test scores for a given testing period, we impose additional sample restrictions for our longitudinal analyses to ensure that comparisons of aggregate measures over time reflect changes in student performance as opposed to changes in the populations of students tested. We have two types of restricted analytic samples.

1. The **2020-21, 2021-22, and 2022-23 school year growth samples** include students who completed and received valid scores for the same benchmark assessment in the same subject, grade level, and district in both the fall and spring of a particular school year (i.e., the 2020-21 growth sample includes students with valid test scores in fall 2020 and spring 2021, while the 2022-23 growth sample includes students with valid test scores in fall 2022 and spring 2023).
2. The **three-year growth sample** includes students with valid test scores in all six semesters between fall 2020 and spring 2023 in the same subject and same district, in the same grade level in both the fall and spring of each school year and progressed to the next consecutive grade level in between school years (e.g., students who were in the 3rd grade in fall 2020 and spring 2021, 4th grade in fall 2021 and spring 2022, and 5th grade in fall 2022 and spring 2023).

To illustrate how these restrictions affect the size and representativeness of the samples in our growth analyses, Table 2.1 shows the total number of districts and students for whom we received spring 2023 student-level data and the subsets of these students who we can and cannot include in each of the more restricted samples. The figures in the top panel represent the exclusions for the 2022-23 school year growth sample. Though not shown here, we use similar exclusions to identify the 2020-21 and 2021-22 school year growth samples for comparisons of fall-to-spring growth for each of the three school years. Figures in the bottom panel show the exclusions for the three-year growth sample.

The full sample for the spring 2023 testing period includes student-level data for 703,693 students from 734 districts[1]. For analyses of 2022-23 growth, we exclude a total of five districts and 24,577 students (about 1% and 3% of all districts and students, respectively). The most common reason why students were excluded from the 2022-23 school year growth sample was that they did not have benchmark assessment data from the fall 2022 semester. The remaining 679,116 students and 729 districts in the 2022-23 school year growth sample participated in comparable benchmark assessments, in the same district, in the same grade levels, for the entire school year.

The three-year growth sample is the most restrictive sample in that it includes only the 295,006 students (from 548 districts) who participated in comparable benchmark assessments in both the fall and spring of the 2020-21, 2021-22, and 2022-23 school years in the same district, in the same grade level in both the fall and spring of the 2020-21 school year, then in the next consecutive grade levels in both semesters of 2021-22 and 2022-23. About 42% of students and 75% of districts in the 2022-23 school year growth sample were also in the three-year growth sample. One major factor driving the high exclusion rate is that more than one hundred districts that chose to use a locally developed benchmark assessment in 2020-21 switched to an assessment from one of the four MDE-approved providers the following year, as the benchmark assessment legislation for 2021-22 allocated new funding for districts to implement

these assessments (2021 PA 48). We include these districts in our year-specific growth analyses but cannot measure growth over a three-year period, as they have incomplete benchmark assessment data. The two districts in the "Different district in prior year(s)" column are charter schools that only serve middle school grades, and hence the students would have had to move into the district from somewhere else after elementary school.

Table 2.1. Spring 2023 Analytic Sample and Reasons for Exclusions from Restricted Samples for Growth Analyses

Exclusions	Districts		Students	
	N	%	N	%
2022-23 School Year Growth Sample and Exclusion Reasons				
Spring 2023 sample	**734**	**100**	**703,693**	**100**
Missing fall 2022 data	-5	-1	-17,630	-3
Different district in fall 2022 than in spring 2023	-0	-0	-6,648	-1
Different test in fall 2022 than in spring 2023	-0	-0	-299	-0
2022-23 school year growth sample	**729**	**99**	**679,116**	**97**
3-Year Growth Sample and Exclusion Reasons				
2022-23 school year growth sample	**729**	**99**	**679,116**	**97**
New K-1 cohorts, not tested in 2020-21	-1	-0	124,903	-18
Missing fall or spring data in prior year(s)	-161	-22	206,213	-29
Different district in prior year(s)	-2	-0	36,111	-5
Different test in prior year(s)	-17	-2	16,883	-2
3-year growth sample	**548**	**75**	**295,006**	**42**

Notes: The counts and percentages in this table do not include data from districts that prepared their own aggregate datasets. The percentages in each column may not total exactly 100 due to rounding.

When possible, we also include data from the 22 districts that prepared their own aggregate files. These aggregate files include benchmark assessment data for another 69,518 Michigan students, meaning that the combined dataset we constructed from both the student-level and district-provided aggregate data represents 773,211 (or about 82%) of all K-8 students in Michigan.

Sample Characteristics and Representativeness

Prior to aggregation, we combined the benchmark assessment data with student characteristics from the Michigan Student Data System (MSDS) Fall General Collection of each school year to identify subgroups and examine the representativeness of each sample based on demographic and academic characteristics. Table 2.2 provides summary statistics of K-8 students in Michigan and by assessment provider.

Table 2.2. Summary Statistics of K-8 Students in All Michigan Districts and by Spring 2023 Assessment						
Demographics (%)	**All MI**	**MAP Growth**	**i-Ready**	**Star 360**	**MDE/ ICA**	**All Vendors**
Female	48.6	48.3	47.6	47.3	48.3	48.7
Asian	3.7	3.2	5.0	1.1	0.2	3.4
Black	18.4	16.3	31.9	6.7	3.9	18.5
Latino	8.9	8.9	9.4	9.5	4.6	9.1
White	62.9	64.2	47.1	78.2	86.7	62.4
Economically disadvantaged	55.8	56.1	58.2	53.4	48.9	56.3
Special education	13.7	13.2	11.8	14.4	12.3	13.0
English learner	7.4	7.7	9.9	2.8	1.6	7.8
N students	947,099	573,904	152,688	60,115	4,830	773,211
% of MI K-8 students	100.0	60.6	16.1	6.3	0.5	81.6

Notes: The "All MI" column includes the full population of K-8 students across Michigan. Each vendor-specific column includes all students who took a MAP Growth, i-Ready, Star 360, or K-2/ICA assessment in spring 2023, respectively. The total number of students in the "All Vendors" column is less than the sum of the four vendor-specific columns because some students took benchmark assessments from more than one vendor.

Overall, the demographic characteristics of students in the full spring 2023 sample are very similar to the statewide population of K-8 students, as is the sub-sample of students who took the MAP Growth assessments. The population of students who took the i-Ready assessment is more racially/ethnically diverse and includes more students from economically disadvantaged households, compared to the full K-8 population[2]. Students who participated in the Star 360 assessments are less diverse than the statewide population, less likely to be economically disadvantaged, and less likely to be eligible for English learner services. Students who participated in DRC assessments (MDE and Smarter Balanced ICAs) are the least likely to be economically disadvantaged or eligible for special education or English learner services.

Table 2.3 presents grade-specific enrollment counts and percentages of enrolled students who are represented in each of our analytic samples. The denominator for each inclusion rate is the aggregate enrollment count across all districts offering a particular benchmark assessment for a particular grade level (e.g., a district may use MAP Growth for some grade levels and a locally developed assessment for others). Since grade-specific enrollment counts and inclusion rates were relatively consistent across our reading and mathematics samples, we provide figures for the percentage of students with valid test scores in at least one subject area. The percentages in this table do not include students from districts that submitted their own aggregate data. These districts reported mathematics and reading outcomes separately but did not indicate how many students participated in benchmark testing for both subjects.

Table 2.3. Percent of Enrolled Students Included in Analytic Samples, All Four Assessment Providers

2020-21 Grade	2021-22 Grade	2022-23 Grade	2022-23 Enrollment	Inclusion Rates by Analytic Sample		
				Spring 2023	2022-23 School Year Growth	3-Year Growth
---	---	K	81,551	93.8	86.5	---
---	K	1st	82,622	94.5	90.4	---
K	1st	2nd	82,860	94.4	90.7	46.9
1st	2nd	3rd	86,520	94.5	91.5	52.7
2nd	3rd	4th	86,807	93.8	91.0	56.5
3rd	4th	5th	88,002	93.8	91.1	59.2
4th	5th	6th	88,122	93.2	89.8	57.6
5th	6th	7th	87,335	92.3	88.7	55.8
6th	7th	8th	88,506	91.5	88.0	56.1

Notes: The "2022-23 Enrollment" column represents the total number of students enrolled in each grade level in 2022-23 in any district that provided student-level benchmark assessment data from the spring 2023 testing period. The remaining columns show the percentages of these students who are represented in the spring 2023, 2022-23 school year, and 3-year growth samples, respectively.

As seen in the table, the spring 2023 sample includes about 92% to 95% of all students who were enrolled in a participating district in 2022-23, depending on the grade level. Slightly fewer students are included in the 2022-23 school year growth sample, as not all students who participated in benchmark assessments in spring 2023 took the same assessment in fall 2022. In comparison, the inclusion rates decrease much more substantially between the 2022-23 school year growth and 3-year growth samples. The largest differences in inclusion rates between the spring 2023, 2022-23 fall-to-spring, and three-year growth samples are for the students who were in kindergarten, first, or second grade in fall 2020. This is likely due to transitions between tests (e.g., students who take the Star Early Literacy assessment in kindergarten, then switch to the Star Math and Reading assessments in first or second grade) or districts that use assessments from different providers for their lower and upper elementary and middle school students.

We incorporated information about districts' instructional modality offerings in 2020-21 and students' learning modalities in 2022-23 into our analytic dataset. We showed in our prior report that most districts offered more than one mode of instruction in 2020-21. On average, students had the option to learn in person for between one-half and two-thirds of the year, while hybrid options were available for about a third of the year and remote instruction was offered throughout the year. Students in districts that used the Star 360, Smarter Balanced ICA, and MDE K-2 assessments tended to have more access to in-person instruction in 2020-21, while i-Ready students had the least access.

The overwhelming majority of students (about 98%) returned to in-person learning by 2022-23. Table 2.4 shows that just under 2% of students received fully remote instruction, while almost none received hybrid instruction or switched between modalities (aside from short-term changes in modality due to a local COVID-19 outbreak, for instance). However, we note that these percentages are based on the 96% of students whose districts reported student-level modality data for the 2022-23 school year; we cannot determine how the remaining 4% of students received their instruction.

Given how few students reportedly received instruction in hybrid or remote formats in the 2022-23 school year, we limit our comparisons to just two subgroups: students who received in-person instruction all year and students who received any other mode of instruction (hybrid, remote, or a combination of modalities). In total, 13,862 students from 168 different districts received remote or hybrid instruction in 2022-23. Seventy percent of these students attended charter schools that have always operated virtually, while the remaining 30% attended traditional public and charter schools that provide face-to-face instruction. Notably, the share of students learning remotely was much higher among students who took the Star 360 assessments (about 9%) than in any other group. We find that this patten is driven by three virtual charter districts: Michigan Virtual Charter Academy, Michigan Great Lakes Virtual Academy, and Highpoint Virtual Academy of Michigan. Across the other 80 districts that used the Star 360 benchmark assessments, about 99% of students received in-person instruction in 2022-23.

Table 2.4. Percent of Students Participating in Each Modality in 2022-23; Overall, by Assessment, and by Access to In-Person Instruction in the Prior School Year

	In-Person	Hybrid	Remote	Multiple	No Data
All students in spring 2023 sample	98.1	0.0	1.9	0.0	3.9
By assessment					
MAP Growth	98.7	0.0	1.3	0.0	4.4
i-Ready	99.1	0.0	0.9	0.0	2.6
Star 360	91.0	0.0	9.0	0.0	1.7
Smarter Balanced ICA/MDE K-2	99.7	0.0	0.3	0.0	22.5
By access to fully in-person instruction in 2020-21					
Never offered	99.4	0.0	0.6	0.0	2.5
Offered for less than half of year	99.3	0.0	0.7	0.0	0.4
Offered for at least half of year	99.7	0.0	0.3	0.0	2.7
Offered all year	99.4	0.0	0.6	0.0	8.2

Notes: Percentages in the first four columns are based on the 96% of students whose districts reported modality information in 2022-23. The percentages in each row may not add to exactly 100% due to rounding. The percentages in the "No Data" column are based on the full population of students with 2022-23 benchmark assessment data.

DATA AGGREGATION AND ANALYSIS

To construct the final aggregate data files used in our analysis, we calculated several indicators of students' academic performance at both the district and state levels, overall and by subgroup[3]. We then combined the resulting district- and state-level aggregate datasets with data from individual districts that prepared their own aggregate data files in lieu of submitting student-level data through the MDH. We completed this process separately for three types of analytic samples to create aggregate measures appropriate for examining student achievement in one specific testing period, growth across a single school year (2020-21, 2021-22, or 2022-23), and longitudinal trends for a consistent group of students across the fall 2020, spring 2021, fall 2021, spring 2022, fall 2022, and spring 2023 testing periods.

The remainder of this section describes each of the indicators of academic achievement and growth, as measured by students' scores on MDE-approved benchmark assessments between fall 2020 and spring 2023, that we construct and analyze to fulfill the reporting requirements in Michigan's benchmark assessment legislation.

Average Achievement Trajectories

The first set of analyses in this report examines trends in average scale across six testing periods: fall 2020, spring 2021, fall 2021, spring 2022, fall 2022, and spring 2023. The MAP Growth, i-Ready, Star 360, and Smarter Balanced ICA benchmark assessments are all scored on vertical scales, which allows us to compare scores for the same group of students on the same assessment across multiple school years as they progress from one grade level to the next. However, because each benchmark assessment has its own unique scale and scale scores are not comparable across assessments (e.g., MAP Growth scores range from 100 to 350 whereas i-Ready scores range from 0 to 800), we present cohort-specific trends in average scale scores separately for each assessment.

As comparison points to help us interpret the overall trends for Michigan students over the six testing periods, we plot them alongside grade-level norms that each assessment provider established before the pandemic. While we use pre-pandemic medians as comparison points for all benchmark assessments, not all providers calculate or present this information in the exact same ways. For instance, although we use the most recent norms that were available for each assessment as of the end of the 2020-21 school year (Curriculum Associates, 2020a; Renaissance Star Assessments, 2021a, 2021b, 2021c; Smarter Balanced Assessment Consortium, 2020b; Thum & Kuhfeld, 2020), the specific year(s) of data, sampling procedures, and methodology to produce norms differ across assessments[4]. We provide additional details about differences in the pre-pandemic norming data and comparison points available for each benchmark assessment in the previous report in this series.

We provide separate results for the MAP Growth, i-Ready, and Star 360 assessments. We omit the DRC assessments from this analysis because the assessments for grades K-2 (the MDE Early Literacy and Mathematics benchmark assessments) are different from the assessments from grades 3-8 (the Smarter Balanced ICAs), which have different scales and are not directly comparable. Given the large number of districts that use only the K-2 assessments from DRC and not the Smarter Balanced ICAs and some districts' use of the Smarter Balanced IAB assessments for some grade levels in some years, there are too few students with Smarter Balanced ICA scores in all three years, so we cannot estimate longitudinal trends for this sample.

Variation in Student Achievement

In addition to analyzing how Michigan students performed on their benchmark assessments on average, we also assess how much their test performance varied. For each benchmark assessment, subject area, grade level, and testing period, we calculate the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of Michigan students' benchmark assessment scores. We compare these values across testing periods to assess how Michigan students' performance has changed over time, both at the middle of the distribution and for lower- and higher-scoring students. We also examine gaps between the scores of Michigan's higher- and lower-performing students and the extent to which this has changed over time.

As pre-pandemic comparison points, we compare the distributions of Michigan students' benchmark assessment scores to the distributions of scores among students in the national norming samples for each assessment. We plot the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of the norming distributions alongside the Michigan-specific distributions for each grade level, subject, and assessment. These comparisons help us to understand how average performance differs between Michigan students and students across the country across the COVID-19 pandemic, and how the amount of variation in student achievement on these assessments differs among these groups.

We provide separate results for the MAP Growth, i-Ready, and Star 360 benchmark assessments. We do not include the DRC assessments in this analysis, as the scale for the MDE K-12 Early Literacy and Mathematics benchmark assessments has changed since the initial (fall 2020) testing period in our analysis, and differences in the populations of students who took the Smarter Balanced ICAs each year make it difficult to make meaningful comparisons about changes in the distribution.

Regression-Adjusted Percentile Ranks

We use the aggregated benchmark assessment data in multiple regression models to estimate relationships between the average achievement in a district over time while controlling for other district characteristics. Multiple regression is a statistical technique that predicts an outcome variable using two or more explanatory variables. This technique estimates the unique relationship between academic achievement in

consecutive semesters and shows how this relationship changed over time, when all else was equal between districts.

The achievement outcomes in each regression model represent the average score on a particular benchmark assessment for a particular district, grade level, and subject at a point in time. We standardize scores relative to the means and standard deviations of student scores from the pre-pandemic norming samples for each assessment to interpret the standardized scores in terms of how a district's average achievement compares to that of students across the country pre-pandemic. For example, a standardized score of negative one indicates that a district's average score is one standard deviation below the national pre-pandemic average.

We estimate the following baseline model:

$$Y_{dgsvt} = \beta_0 + \beta_1 S21_t + \beta_2 F21_t + \beta_3 S22_t + \beta_4 F22 + \beta_5 S23_t + \boldsymbol{\beta_6 DCHAR_{dgt}} + \delta_g + \epsilon_t$$

where Y_{dgsvt} is the average standardized test score of students in district *d*, grade *g,* completing subject test *s* from assessment vendor *v*, in semester *t*. *S21, F21, S22, F22,* and *S23* are binary indicators identifying the semester associated with the outcome of interest, Y_{dgsvt}. The coefficients on these indicators, shown here as β_1 through β_5, describe the change in average standardized test score attributable to each semester, spring 2021, fall 2021, spring 2022, fall 2022, and spring 2023, respectively, relative to fall 2020.

We control for a set of district characteristics, $\boldsymbol{DCHAR_{dgt}}$, including the proportions of students in each district-grade who are female, of different races/ethnicities, economically disadvantaged, eligible for special education services, and who are English learners. We mean-center these characteristics in each school year so that a value of zero represents the state average, allowing us to interpret coefficients for each time period as the predicted achievement for an average district. δ_g is a grade fixed effect which controls for differences in standardized test scores that are unique to a particular grade level, thus enabling us to compare across grades in the same model. To ease interpretation of these values, we convert each standardized test score estimate to a percentile rank that describes where Michigan students' average achievement falls along the pre-pandemic national norming distributions for each assessment. For example, a percentile rank of 50 indicates that Michigan students scored at the pre-pandemic national average.

We estimate additional variations of this model that include interactions between each semester-specific time indicator with subgroup indicators to estimate results separately by race/ethnicity, economically disadvantaged status, and access to in-person instruction during the 2020-21 school year. For our instructional modality analysis, we assign students to each modality type based on the number of months their district offered in-person instruction: zero months, one to four months, five to eight months, or all nine months of the 2020-21 school year.

Given concerns about the reliability of fall 2020 benchmark assessment scores for lower elementary students, we limit our main models to cohorts of students who were in the 3rd, 4th, 5th, or 6th grade in 2020-21. However, we fit an additional model that includes students who were in kindergarten, 1st, or 2nd grade in 2020-21. This model includes interactions between each grade-level indicator and the semester-specific time indicators, allowing us to estimate separate trends for each individual grade level. We estimate separate variations of these models for the samples of districts that used the MAP Growth and i-Ready assessments, as well as a combined model with both groups of districts. In the combined model, we include a "vendor" fixed effect to account for differences between these tests. We do not include the Star or Smarter Balanced samples due to low counts of students and districts within some of the grade levels and subgroups of interest for these analyses.

Proficiency Rates

For a general understanding of how Michigan students' performance on benchmark assessments compares to state standards for grade-level proficiency, we use information from each benchmark assessment provider to map students' benchmark assessment scores to approximate M-STEP proficiency levels.

NWEA, Curriculum Associates, and Renaissance Learning each developed their own crosswalks between scale scores on their benchmark assessments and M-STEP proficiency levels using an equipercentile linking method (Curriculum Associates, 2020b; NWEA, 2020; Renaissance Learning, 2019). This means that, for a group of students who took both the M-STEP and a particular benchmark assessment, they identified score cut-offs for their benchmark assessments so that the percentage of students in each proficiency category would perfectly match the percentages of students who scored within the same category on the M-STEP (e.g., if 20% of students in this sample scored in the "advanced" level on the M-STEP, the benchmark assessment vendors would have set their cut-off so that exactly 20% of students fall within their "advanced" category as well). This process was not necessary for the Smarter Balanced assessments, as DRC designed both the M-STEP and Smarter Balanced Assessments and derives the scores for both of these assessments from the same underlying scale (MDE, 2019; Smarter Balanced Assessment Consortium, 2020a, 2021) making it possible to simply convert the M-STEP score cut-offs for each proficiency category to Smarter Balanced scale scores.

After using these crosswalks to map students' benchmark assessment scores to equivalent M-STEP proficiency levels, we compare these M-STEP proficiency equivalencies to the actual M-STEP proficiency rates for students in the same districts in 2018-19. This provides a reference point to help gauge how Michigan students' achievement between fall 2020 and spring 2023 differs from the achievement of students who attended the same districts in 2018-19.

Student Growth

Although we can compare average scale scores and regression-adjusted trends across grades, it is important to note that the "typical" amount of test score growth over the course of a school year often differs by grade level, subject, and initial achievement level. To account for these differences, we compare changes in students' scale scores between the fall and spring of each school year to pre-determined norms for "typical growth" on a particular assessment, subject area, and grade level for students who scored within the same range on their fall assessment.

The growth norms for each assessment are defined in slightly different ways and have slightly different meanings. For students who completed MAP Growth assessments, we use as a growth norm the 50th percentile of the fall-to-spring conditional growth distribution for students with the same initial percentile rank (Thum & Kuhfeld, 2020). For Curriculum Associates, we use typical growth targets from the i-Ready assessment growth models which indicate the median growth of students in the same grade level with the same initial placement levels nationwide pre-pandemic (Curriculum Associates, 2020a). For Star 360 and Smarter Balanced ICA assessments, we use pre-pandemic scale score distributions to identify "typical growth" as the change in scale scores necessary for a student to maintain the same percentile rank in the spring as they did in the fall (Renaissance Star Assessments, 2021a, 2021b, 2021c; Smarter Balanced Assessment Consortium, 2020b). These measures represent the increase in scale scores necessary for a Star 360 or Smarter Balanced ICA student who scored, for example, in the 25th percentile in fall 2020 to also score in the 25th percentile on their spring 2021 benchmark assessment.

While these growth thresholds help us gain a better understanding of academic growth among Michigan students during the COVID-19 pandemic, it is important to note that we are using summary tables released by each assessment provider to assign growth norms to groups of students; each assessment provider uses sophisticated student-level models to derive growth measures and we are unable to perfectly replicate those measures from just the summary tables and the aggregate district-level data made available under the Return to Learn legislation. For example, most assessment providers account in their growth calculations for the number of instructional days a student received between two testing occasions, based on the test dates relative to the district's instructional calendar. For our aggregate, statewide analyses, we cannot account for the exact amount of instructional time between each student's annual fall and spring assessments and accordingly, we assign growth norms as though the timing were the same for all students.

To assess students' actual growth relative to "typical growth," we first calculate the difference between each student's spring and fall scale scores from the same school year, then compare this fall-to-spring change to the appropriate growth norm (i.e., the *typical scale score increase* based on the assessment provider, grade level, subject, and the student's initial achievement level). Before aggregating the data to the district level, we group students into three categories that describe their fall-to-spring growth for each school year: students who did not demonstrate any growth at all (i.e., their scale scores remained the same or decreased from fall to spring); students who achieved partial growth (i.e., their scale scores increased from fall to spring, but the increase was less than the typical growth for their grade, subject, and initial achievement level); and students who met or exceeded their growth targets (i.e., their scale scores increased by an amount equal to or greater than the typical growth for their grade, subject, and initial achievement level). We examine patterns in the percentages of students in each of these categories each year and by subject overall and by assessment vendor, grade level, demographic group, and mode of instruction.

PURPOSE AND FRAMING OF SUBGROUP COMPARISONS

The legislation requires this report to disaggregate student achievement and growth data by demographic group (2020 PA 149, 2021 PA 48, 2023 PA 103). To fulfill this requirement, we show comparisons of regression-adjusted percentile ranks and student growth outcomes across different populations of students, including race/ethnicity subgroups. Recently, NWEA researchers published the statement below, which mirrors EPIC's own beliefs about the race/ethnicity analyses:

> *We recognize that focusing on differences across race and ethnicity groups can have negative implications, as it can perpetuate a deficit-oriented perspective that blames students and fails to recognize academic strengths, which may not be accurately reflected in standardized metrics. At the same time, it is crucial to disaggregate outcomes by race and ethnicity to shine light on the profound inequities existing within our education system. Those inequities were stark before the pandemic and have only widened dramatically over the last three years.*
>
> *In this context, we share data on which students were disproportionately harmed during the pandemic, not to assign blame but to highlight the students to whom we owe, as Gloria Ladson-Billings coined, the greatest "educational debt" (Ladson-Billings, 2006). These data underscore the scope of the resources and supports schools must provide to address the cumulative impacts of the pandemic and rectify the harm these students have experienced. (Lewis & Kuhfeld, 2023)*

SUMMARY

The analyses included in this report are based on data representing 90% of districts (769 of the 852 total districts) and 82% of students in the state (773,211 of the 947,099 total students). However, those who are represented in our analyses may not be reflective of those who are not included. Moreover, some of our analyses are limited to certain subsets of the full dataset, depending on the type of data needed for the analysis and the information available about each assessment. Table 2.5 summarizes the specific samples, grade levels, and assessments that are represented in each of our analyses.

Table 2.5. Samples, Grade Levels, and Assessments Included in Each Analysis

Analysis	Sample	Grades	Assessments	Notes
Average Achievement Trajectories	3-year growth	K-8	MAP Growth, i-Ready, Star 360	Grades 1-8 only for Star Math
Variation in Student Achievement	Full sample	K-8	MAP Growth, i-Ready, Star 360	Grades 1-8 only for Star Math
Regression-Adjusted Percentile Ranks	3-year growth	3-8	MAP Growth & i-Ready	Some models also include K-2
Proficiency Rates	Full sample	3-7	All	M-STEP grade levels only
Student Growth	School year growth	K-8	All	All grades/vendors with growth norms

Notes: The average achievement trajectory and regression-adjusted percentile rank analyses include district-provided aggregate data. We do not include kindergarten scores for Star Math because the assessment is only normed for students in 1st grade and above. The exact grade levels included in our student growth analysis vary across vendors, as we can only include grade levels for which the vendor has growth norms available.

While it is important to keep in mind the limitations of the data when interpreting results, the report nonetheless helps deepen our understanding of how Michigan public school students progressed academically between the fall 2020 and spring 2023 semesters. The analyses presented in Section Three continue to expand on the descriptive results presented in our previous reports, providing a more refined estimate of academic growth by incorporating another year of assessment data and comparing academic trajectories of Michigan students to pre-pandemic trajectories of students from across the country.

Section Three: Results

In this section, we summarize outcomes for Michigan students on benchmark assessments that districts administered between fall 2020 and spring 2023. We show changes in average scale scores as well as regression-adjusted percentile ranks using multiple regression models that control for district and grade-specific student characteristics. We also assess how Michigan student test score growth in 2020-21, 2021-22, and 2022-23 compares to pre-determined growth norms that each assessment provider established before the COVID-19 pandemic, and how these patterns vary across students from different demographic groups and instructional modalities.

It is important to remember that different districts use different benchmark assessments. While the MDE-approved assessments all measure similar constructs (e.g., math or reading achievement), there are slight differences in the way they were designed, their intended purposes and the content they cover. Moreover, each assessment has its own unique scale, and scores are not comparable across assessments. For these reasons, we conduct certain analyses separately by vendor. Where possible, we include combined analyses that rely on standardized metrics that have similar meanings across assessments. These combined analyses help us to understand what the results from the separate assessments collectively tell us about student learning and recovery for the state as a whole.

AVERAGE ACHIEVEMENT TRAJECTORIES

Average scores on vertically scaled assessments typically increase over time and across grade levels, with slight decreases during the summer months (e.g., see McEachin & Atteberry, 2017; Quinn et al., 2016). As students become older and advance to higher grade levels, year-to-year increases in average scores typically become smaller and gaps between higher- and lower-achieving students become larger (von Hippel & Hamrock, 2019). However, students' learning experiences during and in the wake of the COVID-19 pandemic have been far from "typical." To understand how Michigan student achievement trajectories between fall 2020 and spring 2023 differ from past norms, we examine trends in average scale scores among the same sample of students across the spring and fall testing periods of all three years.

In previous reports, we showed that average scores for students in most grade levels were close to pre-pandemic national norms in the fall of the 2020-21 school year. As a notable exception to this pattern, fall 2020 scores for students in lower elementary

grades were often substantially *above* pre-pandemic norms, which we attribute mostly to more favorable at-home testing conditions for younger students rather than a true reflection of their achievement at that time. In all grade levels, average scores increased over the course of the year, but they did so at a slower rate than those in the pre-pandemic norming samples and, as a result, students fell further below norms by the end of the year. Scores for upper elementary and middle school students generally remained below pre-pandemic norms throughout the 2021-22 school year, while trends for students in lower elementary grades varied depending on the assessment and subject area.

HOW TO INTERPRET FIGURES 3.1-3.3

In these figures, **green and blue points** represent average math and reading scale scores, respectively, in each testing period.

Average scores for the same cohort of students are connected by **solid lines**, showing the cohort's achievement trajectory over time. Each point is labeled to show what grade level students were in during a particular testing period. **Dashed lines** represent pre-pandemic norms for a given assessment, subject area, and grade level. The **shaded area between pairs of solid and dashed lines** shows the difference between the average score for Michigan students and the pre-pandemic norm.

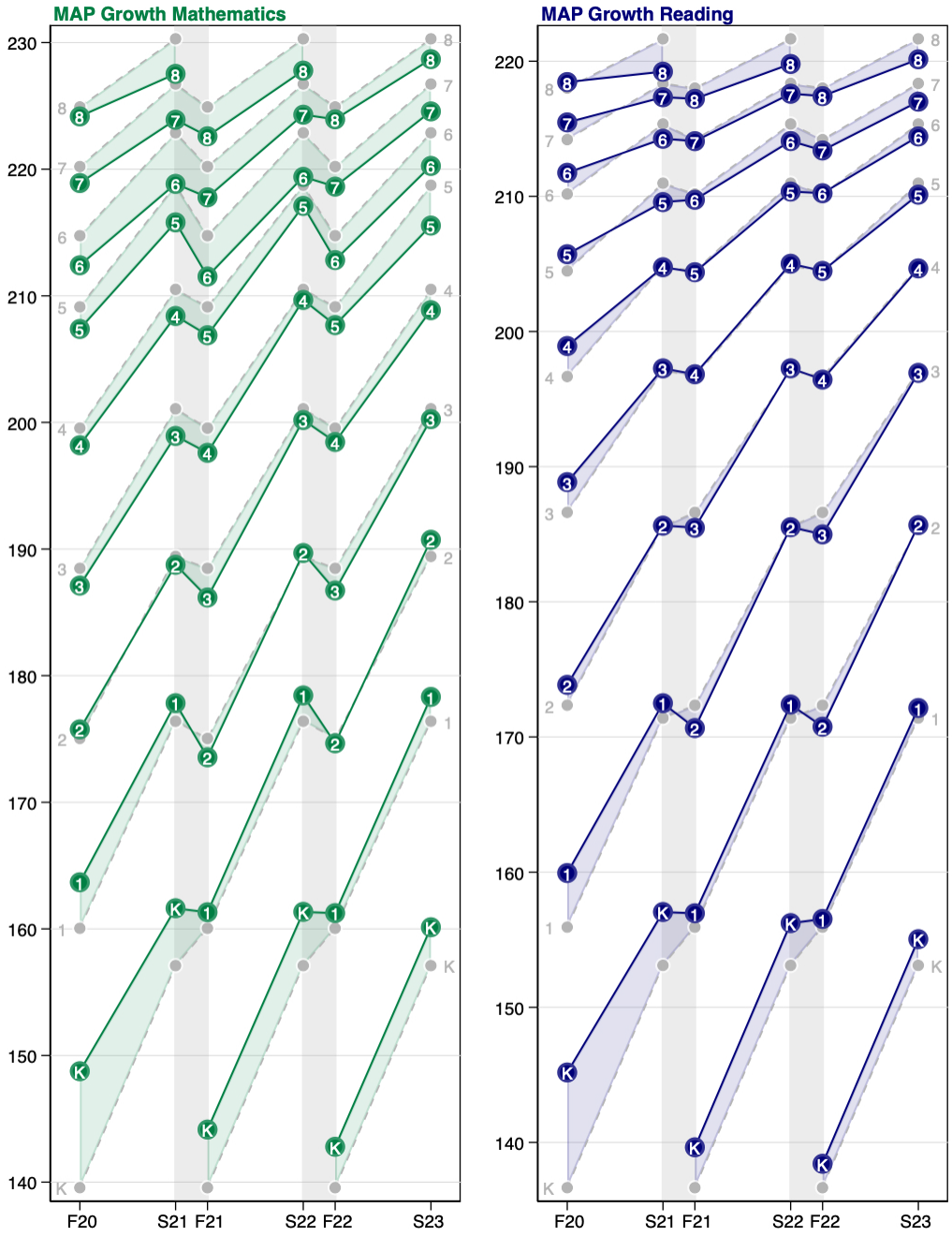
The **y-axis scales** in each figure extend approximately from the kindergarten fall norm for each assessment to the 8th grade spring norm. Although the exact numbers differ slightly between subjects and differ greatly across vendors, the total distance from the bottom to the top of each y-axis always represents the range of grade-level norms from the beginning of kindergarten to the end of 8th grade.

Figures 3.1 to 3.3 extend the trend analyses from our past reports to include new data from the fall 2022 and spring 2023 testing periods. Average scores for most vendors, grade levels, and subjects remained below national norms by the end of the 2022-23 school year. Gaps between average scores and pre-pandemic norms remained stagnant for upper elementary students over the course of the year, suggesting that these students learned at similar rates as the pre-pandemic norming samples. While this means that, on average, these students did not fall any further below national norms than they were at the beginning of the year, students would have had to experience a more accelerated rate of learning (i.e., faster rates than students in the pre-pandemic norming samples) for their scores to "catch up" to pre-pandemic levels. Middle school students, on the other hand, generally fell somewhat further below national norms between the fall and spring testing periods.

However, there were also some exceptions where average scores approached or even exceeded pre-pandemic norms in 2022-23. In districts that used the MAP Growth assessments (shown in Figure 3.1), kindergarten through 2nd grade students reached or exceeded pre-pandemic math norms in both the fall 2022 and spring 2023 testing periods. In these same districts, kindergarten, 1st, and 4th grade students remained at or slightly above pre-pandemic reading norms throughout the 2022-23 school year, while 2nd and 3rd graders started the year below norms but experienced an accelerated learning rate throughout the year and were able to reach pre-pandemic achievement levels by spring 2023. Figure 3.2 shows that middle school students in i-Ready districts, on average, scored at or above pre-pandemic norms all year in math[5]. These students started the year above norms in reading but decreased relative to national norms by the end of the year. Elementary students who took the Star assessments maintained average math scores above national norms throughout the 2022-23 school year (shown in Figure 3.3), and although their fall reading scores were below pre-pandemic norms, all except the 1st grade cohort experienced accelerated learning rates and ended the year closer to national reading norms than where they started[6].

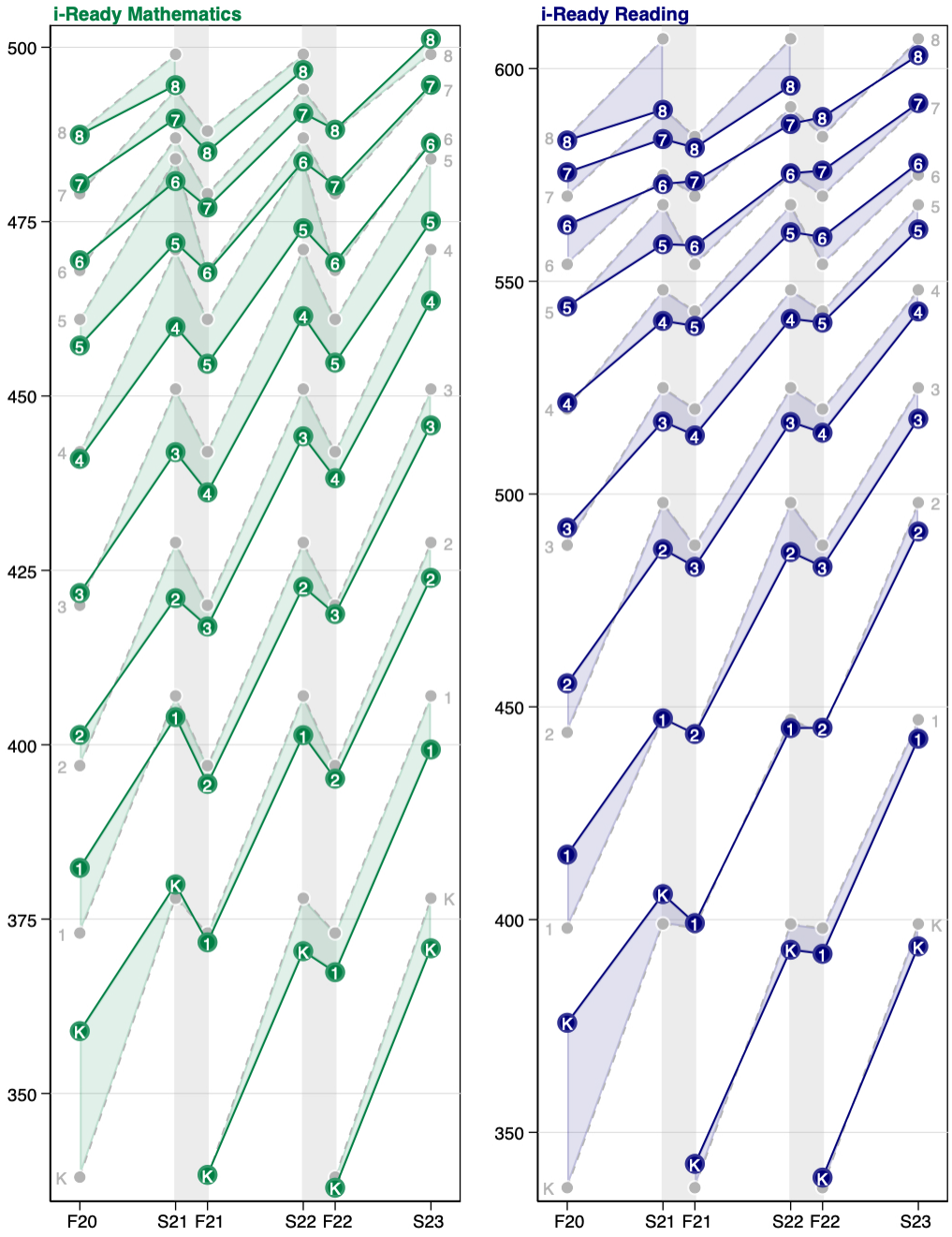
It is important to note that some differences we observe in achievement trends across vendors or grade levels may be driven by differences in the populations of students and districts that participated in each benchmark assessment. For instance, the i-Ready sample consists primarily of students from urban districts that operated remotely in 2020-21, whereas the Star 360 sample consists primarily of smaller districts in suburbs, towns, and rural areas that were more likely to stay in-person in 2020-21 (see Section Two of this report). While we generally see larger discrepancies between Michigan students' scores and pre-pandemic norms in middle school grades, this is likely due in part to the fact that student achievement on vertically scaled tests typically varies more in upper grade levels in general. In some cases, these patterns may also reflect differences in the populations of students in different grade levels who take these assessments. For instance, as we noted in Section Two, about 9% of students in the Star 360 sample attended virtual charter schools, which primarily serve students in higher grade levels. Thus, differences in the characteristics of students who attend virtual charter schools may affect achievement trends on the Star 360 assessments to a greater degree in middle school grade levels than in elementary grades. It is also possible that some districts administer the Star 360 assessments to all their elementary students but only as a progress monitoring tool for middle students who are receiving certain interventions, which could result in very different achievement trends across grade levels.

Figure 3.1. Trends in Average Scale Scores, MAP Growth RIT Scale



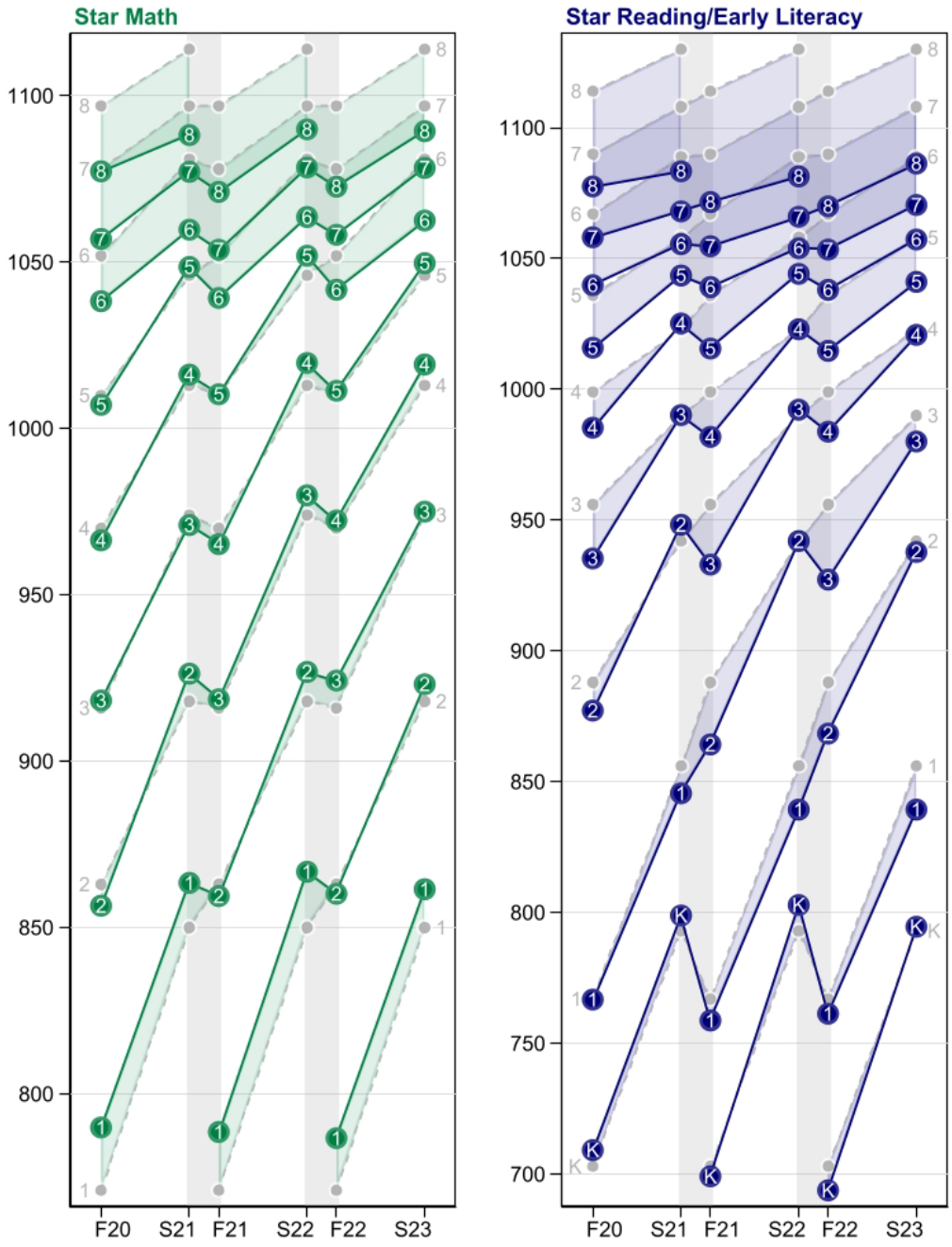
Notes: Dashed lines represent pre-pandemic national norms. Averages only include students with benchmark assessment scores for every possible testing period. The y-axis scales range from the kindergarten fall norm to the 8th grade spring norm, which differ slightly across subjects.

Figure 3.2. Trends in Average Scale Scores, i-Ready



Notes: Dashed lines represent the 2018-19 median for MI students. Averages only include students with benchmark assessment scores for every possible testing period. The y-axis scales range from the kindergarten fall norm to the 8th grade spring norm, which differ slightly across subjects.

Figure 3.3. Trends in Average Scale Scores, Star 360 Unified Scale



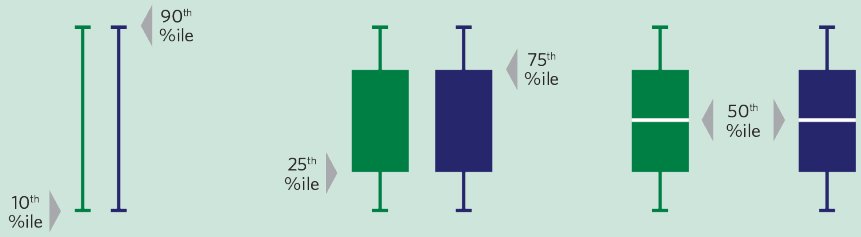
Notes: Dashed lines represent pre-pandemic national norms. Averages only include students with benchmark assessment scores for every possible testing period. The y-axis scales range from the kindergarten fall norm to the 8th grade spring norm, which differ slightly across subjects.

VARIATION IN STUDENT ACHIEVEMENT

While Figures 3.1 to 3.3 help us understand how students performed *on average*, these trends do not necessarily reflect all students' learning experiences or achievement outcomes. To better understand how student achievement trends varied across students, we examine the distribution of scale scores in each testing period. We show distributions for each of the spring testing periods in Figures 3.4 to 3.6 and also provide fall score distributions in Appendix Tables A.2.1 to A.2.3.

HOW TO INTERPRET FIGURES 3.4 TO 3.6

These figures show the **distribution of benchmark assessment scores** for students in each grade level in the spring of each school year. We use lighter shades of green and blue to depict the distribution of scores for the **pre-pandemic norming sample**, and darker shades of the same colors to depict the distributions of **Michigan students'** benchmark assessment scores.



Each vertical line shows the **10th and 90th percentiles** of students' spring scores.

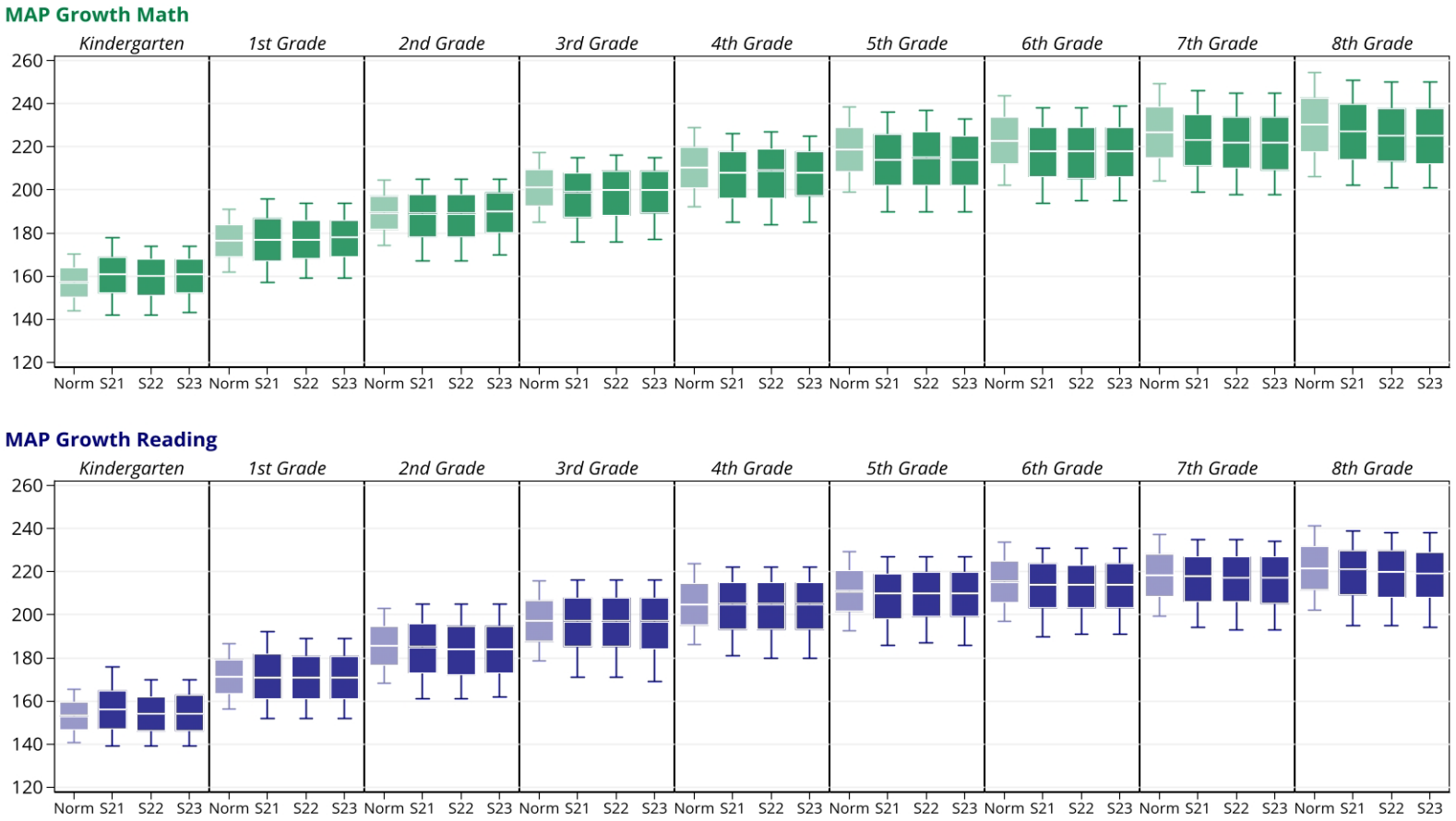
While each rectangle shows the **25th and 75th percentiles**.

The horizontal white line near the center of each rectangle represents the **50th percentile or median**.

We use a "box-and-whisker" design to show not only the average achievement but also how achievement varied across a group of students. Comparing the **positions** of boxes and whiskers across groups of students tells us how **achievement levels** differed between the groups. Comparing the total **lengths** of boxes and whiskers across groups of students tells us how the **variation in student achievement** differs between the groups.

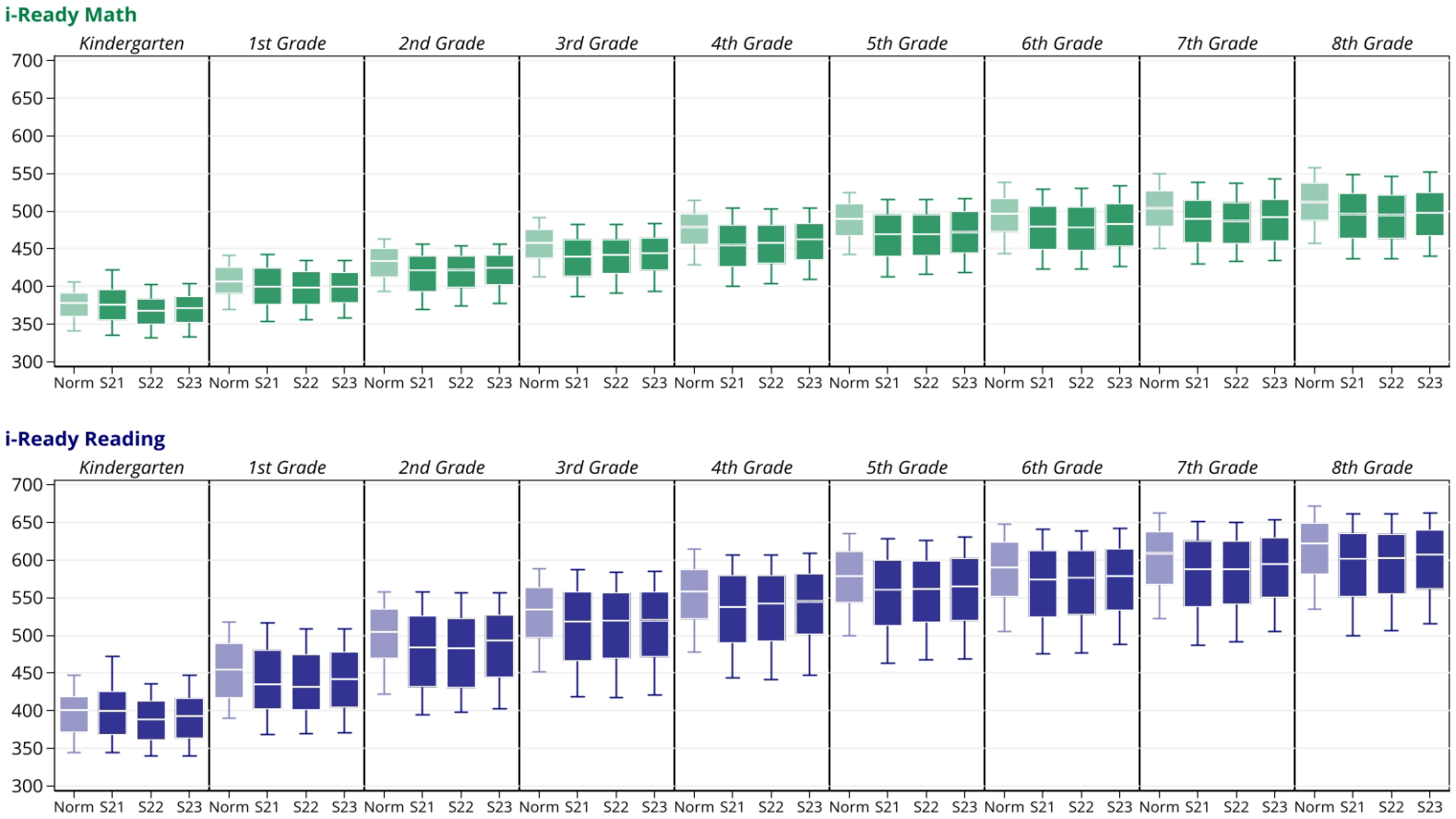
As we would typically expect to see on a vertically scaled test, the centers of these distributions (both for the national norming samples and for Michigan students) move upwards across grade levels, indicating that student scores increase as they become older and reach higher grade levels. The distributions span wider ranges of scale scores in upper grade levels, indicating that there is more variation in student achievement at higher grade levels.

Figure 3.4. Distribution of Spring Scale Scores by School Year and Grade Level, MAP Growth



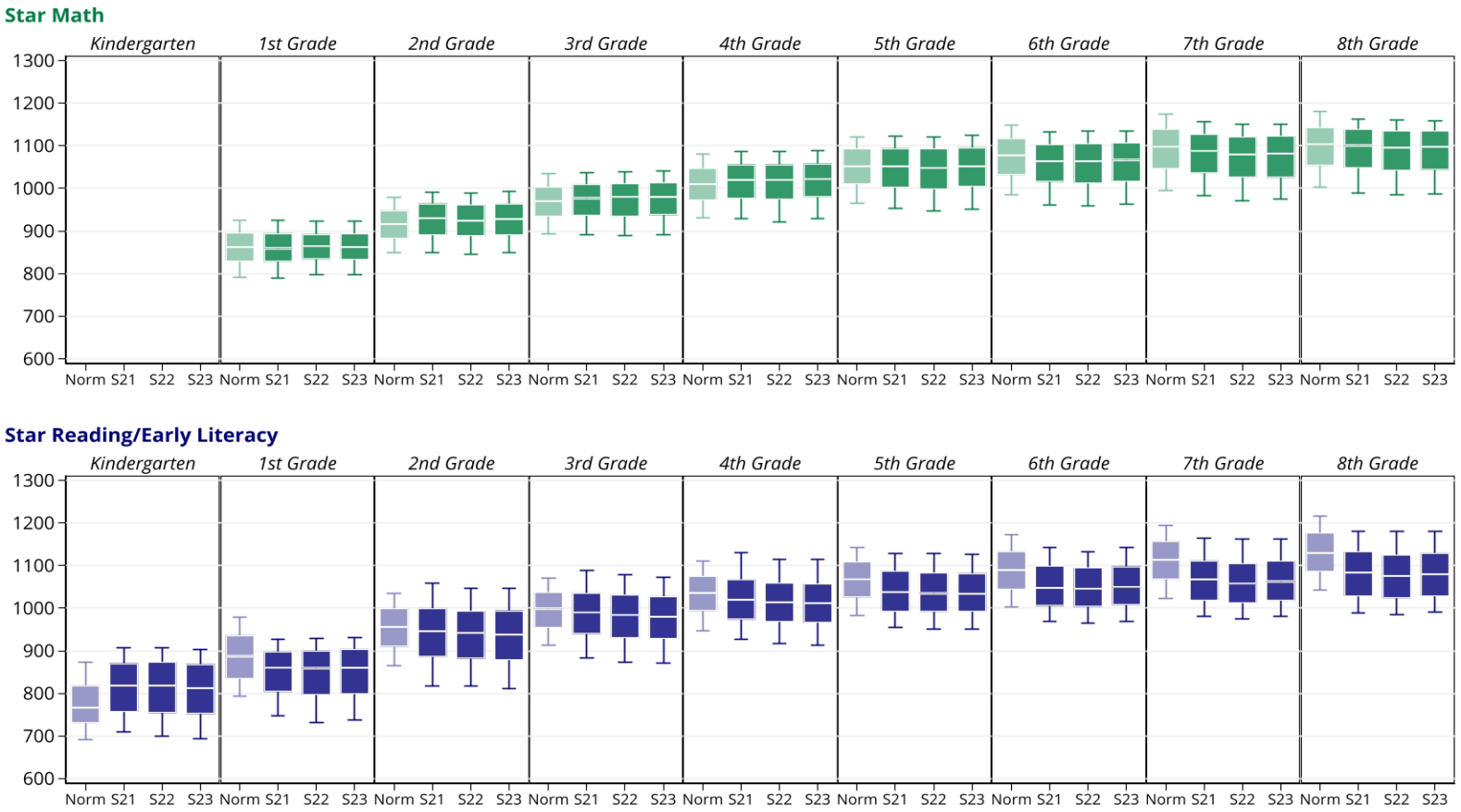
Notes: Each vertical line shows the 10th and 90th percentiles of students' spring scores, while each rectangle shows the 25th, 50th, and 75th percentiles. "Norm," "S21," "S22," and "S23" represent the pre-pandemic national norm and the spring 2021, spring 2022, and spring 2023 testing periods, respectively.

Figure 3.5. Distribution of Spring Scale Scores by School Year and Grade Level, i-Ready



Notes: Each vertical line shows the 10th and 90th percentiles of students' spring scores, while each rectangle shows the 25th, 50th, and 75th percentiles. "Norm," "S21," "S22," and "S23" represent the pre-pandemic national norm and the spring 2021, spring 2022, and spring 2023 testing periods, respectively.

Figure 3.6. Distribution of Spring Scale Scores by School Year and Grade Level, Star 360



Notes: Each vertical line shows the 10th and 90th percentiles of students' spring scores, while each rectangle shows the 25th, 50th, and 75th percentiles. "Norm," "S21," "S22," and "S23" represent the pre-pandemic national norm and the spring 2021, spring 2022, and spring 2023 testing periods, respectively.

Consistent with our findings from Figures 3.1 to 3.3, the centers of the distributions for Michigan students in 2020-21 are generally lower than the centers of the norming distributions for the same grade levels, indicating that the median scores for Michigan students are lower than pre-pandemic national medians. In most cases, the Michigan distributions also span a wider range of scale scores than the norming distribution. This suggests that the gaps between Michigan's higher- and lower-achieving students are larger than pre-pandemic national norms.

While spring 2021 scores for lower elementary students did not always shift below the national median, their distributions widened more than those for higher grade levels, especially among the top 25% of Michigan students. This suggests that, even at the end of the 2020-21 school year, many early elementary students received unusually high scores on their benchmark assessments. Considering that these same patterns are most stark for i-Ready districts (which were more likely to operate remotely all year) and even more pronounced and consistent across vendors in students' fall 2020 scores (shown in Appendix Figures A.2.2 to A.2.3), this is likely due to more favorable conditions in at-home testing.

In most grade levels, median scores changed relatively little across the 2020-21, 2021-22, and 2022-23 cohorts and distributions remained wider than pre-pandemic norms. On most assessments, the lowest-scoring Michigan students (those in the bottom 10%) had far lower achievement than the bottom 10% of students in the national norming samples. The highest-scoring students (top 10%), on the other hand, scored similarly to or slightly lower than the top 10% of the norming sample. These patterns suggest that disruptions to student learning and instruction during the COVID-19 pandemic had a disproportionate effect on lower-achieving students. In i-Ready districts, where students were least likely to have had access to in-person instruction in 2020-21, we find some evidence of learning recovery for students at the bottom of the distribution, particularly in reading at the middle school level.

REGRESSION-ADJUSTED PERCENTILE RANKS

The trends in Figures 3.1 to 3.6 help us to understand how Michigan students in each grade level who took a given assessment performed on average and the extent to which their performance varied. However, it is difficult to discern what these separate trends for each grade and assessment vendor mean for the population of Michigan students overall, given the vast differences in the types of districts and students who participated in each assessment (shown in Table 2.2) and differences in achievement norms across grade levels. To assess the performance of Michigan students overall, we use a regression analysis approach that controls for differences between grade levels, assessment vendors, and demographic characteristics of students in each

district (i.e., gender, race/ethnicity, economically disadvantaged status, special education status, and English learner status). We show the results from these analyses in Figures 3.7 to 3.10. We also provide separate results for each assessment vendor and grade level in Appendix Tables A.3.1 to A.3.2.

HOW TO INTERPRET FIGURES 3.7 TO 3.10

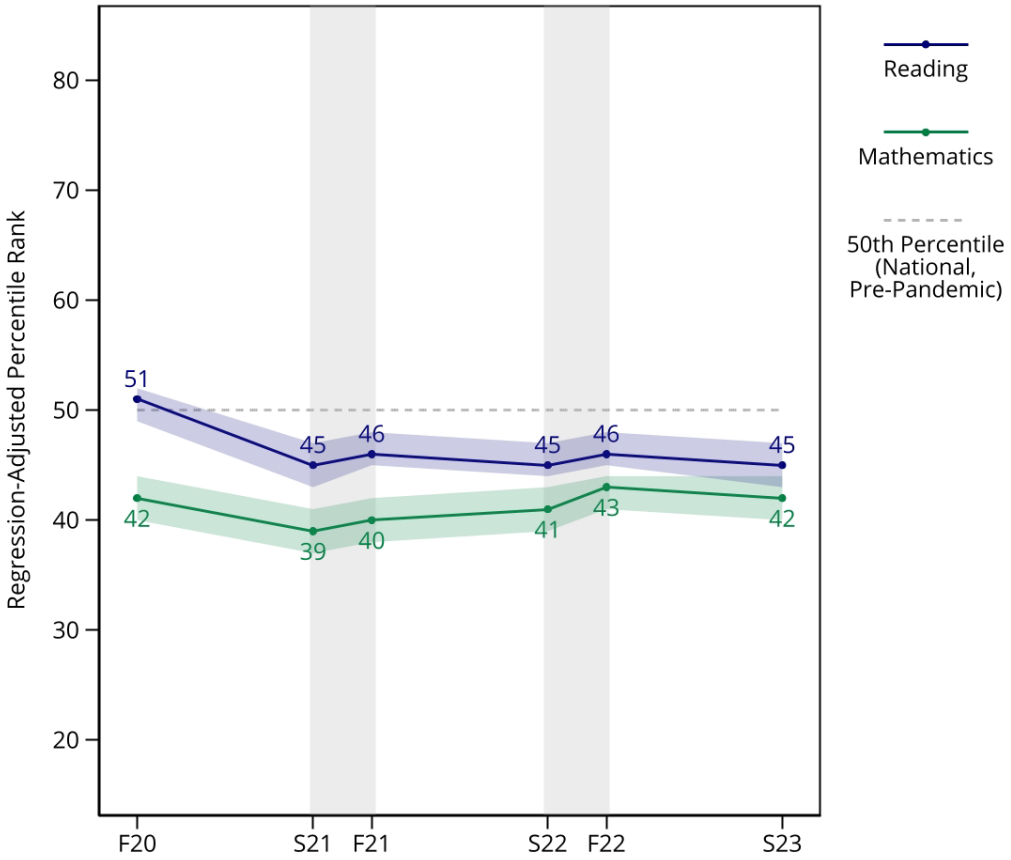
These figures show adjusted trends in average math and reading scores, standardized relative to pre-pandemic national norms for each vendor, grade, subject, and testing period. To ease interpretation of these values, we convert each estimate to a **percentile rank** to show where Michigan students' average scores fall relative to the national norming distributions for each assessment.

A percentile rank of 50 indicates that Michigan students scored at the pre-pandemic national average. If students learned at a rate consistent with pre-pandemic norms, we would see a **flat trend line**, indicating that Michigan students maintained the same percentile rank over time. If students learned at a slower rate than the norming sample, we would see a **decreasing trend**. If students' relative achievement decreased during the pandemic, **they would need to learn at a *faster* rate than the norming sample** to achieve the same percentile ranks they did before the pandemic.

The **shaded areas above and below each trend line** show the 95% confidence interval for each percentile rank estimate. This represents the range of values that the "true" percentile rank for Michigan students is likely to fall within, given that our estimates are based on a sample of students and not the full population. If the shaded area overlaps with the grey dashed line, this means that the estimate is not significantly different from the pre-pandemic national average.

While the data for this study does not include test scores from before the 2020-21 school year, evidence from Michigan's state summative assessments suggests that student learning rates had already declined to some extent in between the initial school closures in March of 2020 and the time that students took their fall 2020 benchmark assessments, especially in math (Strunk et al., 2023). We find that, on average, students began the 2020-21 school year slightly above pre-pandemic norms in reading, and below national norms in math. As Figure 3.7 shows, average MAP Growth and i-Ready scores for 3rd-6th grade students in fall 2020 were at the 51st percentile in reading and in the 42nd percentile in math, relative to students across the country who took the same assessments pre-pandemic.

Figure 3.7. Regression-Adjusted Percentile Ranks by Semester (MAP Growth & i-Ready, 3rd-8th Grade)



Note: These regression estimates are based on district-grade-average scores across students with MAP Growth or i-Ready scores in every possible testing period. We standardized scores relative to pre-pandemic national norms for each assessment and converted all estimates into percentile ranks. Models control for student demographics.

By spring 2021, students fell further below norms in both subjects. These declines were steeper in reading than in math, however, given that math scores were already significantly below national norms at the beginning of the year, spring 2021 percentile ranks were still lower in math than in reading despite the steeper declines in reading that year. In 2021-22 and 2022-23, average reading scores remained at about the same percentile rank, suggesting that students learned at about the same rate as the pre-pandemic norming sample, but not at the accelerated rate that would have been necessary to return to their fall 2020 percentile ranks. Math achievement, on the other hand, increased slightly after 2020-21. By the end of 2022-23, Michigan students' average math achievement reached the same percentile rank as in fall 2020. However, this does not mean that they have recovered to pre-pandemic levels, considering that

their relative math achievement was likely higher than this before the initial disruptions to student learning at the end of the 2019-20 school year.

We use a similar approach to estimate gaps in average achievement across subgroups of students after controlling for differences between tests, grade levels, and the composition of students within a district. Trends in regression-adjusted percentile ranks for students who are economically disadvantaged and their more advantaged peers (shown in Figure 3.8) generally resemble the overall trends in Figure 3.7, with slight improvements over time in math and little change in reading in 2021-22 and 2022-23. However, we find a few notable differences across trends for students of different races/ethnicities. Figure 3.9 shows that the initial achievement declines in 2020-21 were much steeper for Black students in both math and reading than for their White or Latino peers. However, Black students also experienced more learning recovery in subsequent school years.

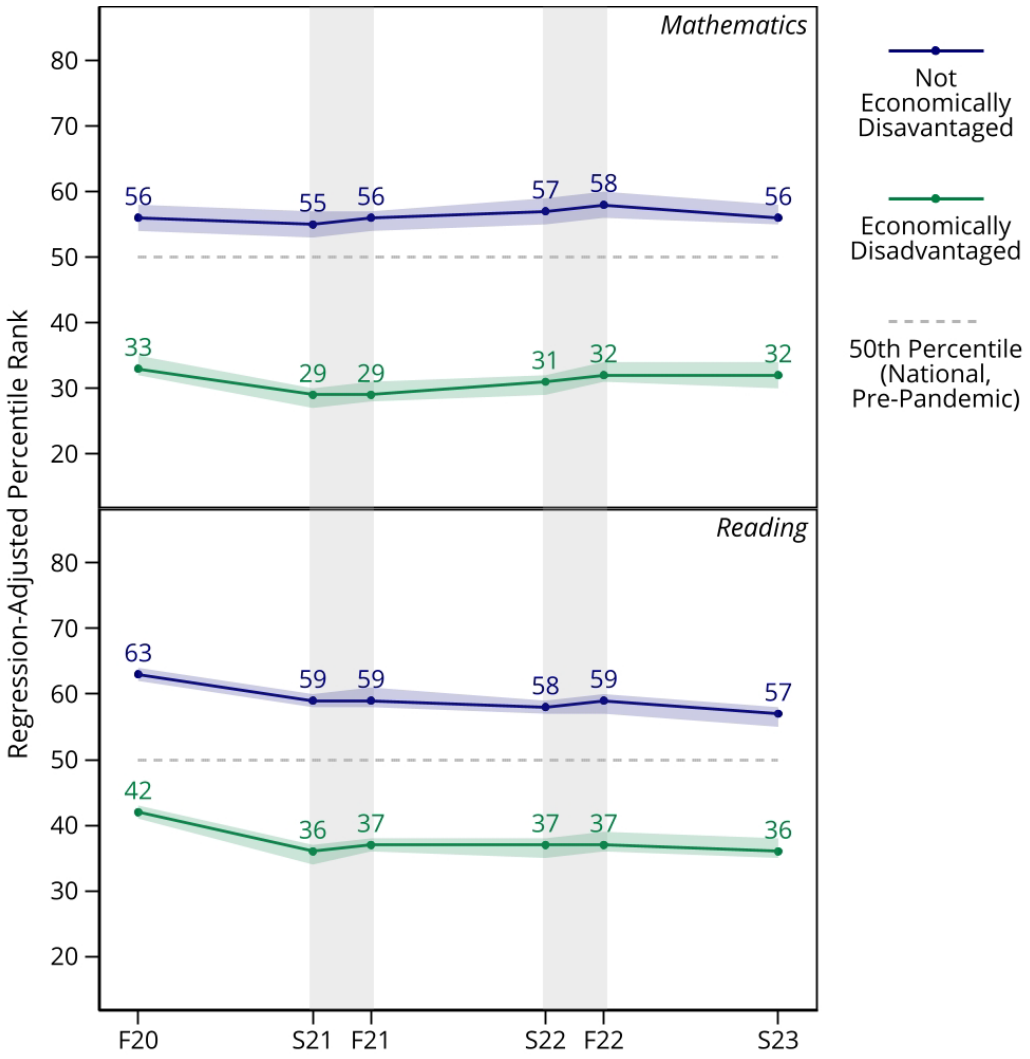
As we showed in past reports, racial and socioeconomic achievement gaps widened slightly in 2020-21. While Figure 3.8 and Figure 3.9 show that these gaps have remained relatively consistent after that year, we find some evidence of slight improvements in 2022-23. For example, average reading scores for Black students were at the 35th percentile of the national distribution in fall 2020, compared to the 55th percentile for White students, a 20 percentile-point gap. This increased to a 24 percentile-point gap by spring 2021 and remained about the same size throughout the 2021-22 school year, eventually reducing to a 19 percentile-point gap in 2022-23.

We find that student learning trajectories differed depending on their access to in-person instruction during the 2020-21 school year. Figure 3.10 shows that average math achievement decreased for students in all districts *except* those that offered in-person instruction in all nine months (September through May) of the 2020-21 school year. While all other districts experienced declines, the extent of these declines varied depending on how much in-person instruction the district offered during the year. Declines were smallest in districts that offered in-person instruction for more than half (5-8 months) of the year, followed by those that were in-person for less than half (1-4 months) of the year, and steepest in districts that did not offer in-person instruction at all. Initial declines in reading achievement follow a similar pattern, except that districts that offered in-person instruction all year also experienced declines, albeit to a lesser extent than other districts.

As a result of these differential declines in 2020-21, gaps in achievement between districts by instructional modality became much larger. These gaps improved to some extent in subsequent school years, especially in reading. For example, average reading scores for districts that were in-person all year and those that did not offer in-person

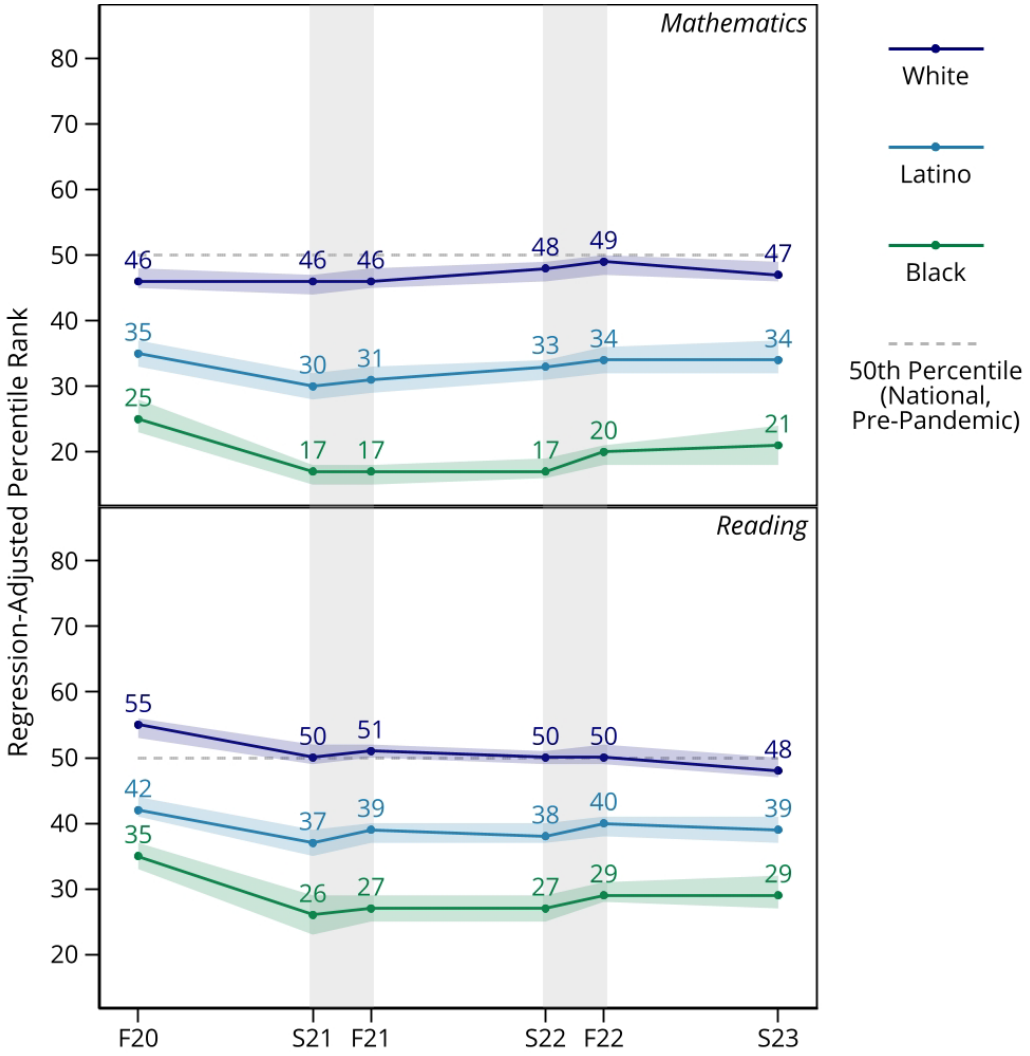
instruction at all were at the 54th and 43rd percentiles of the national norming distributions in fall 2020, an 11 percentile-point gap. This increased to a 17 percentile-point gap by spring 2021, reducing to 13 percentile points in spring 2022 and returning to 11 percentile points in spring 2023.

Figure 3.8. Regression-Adjusted Percentile Ranks by Semester and Economically Disadvantaged Status (MAP Growth & i-Ready, 3rd-8th Grade)



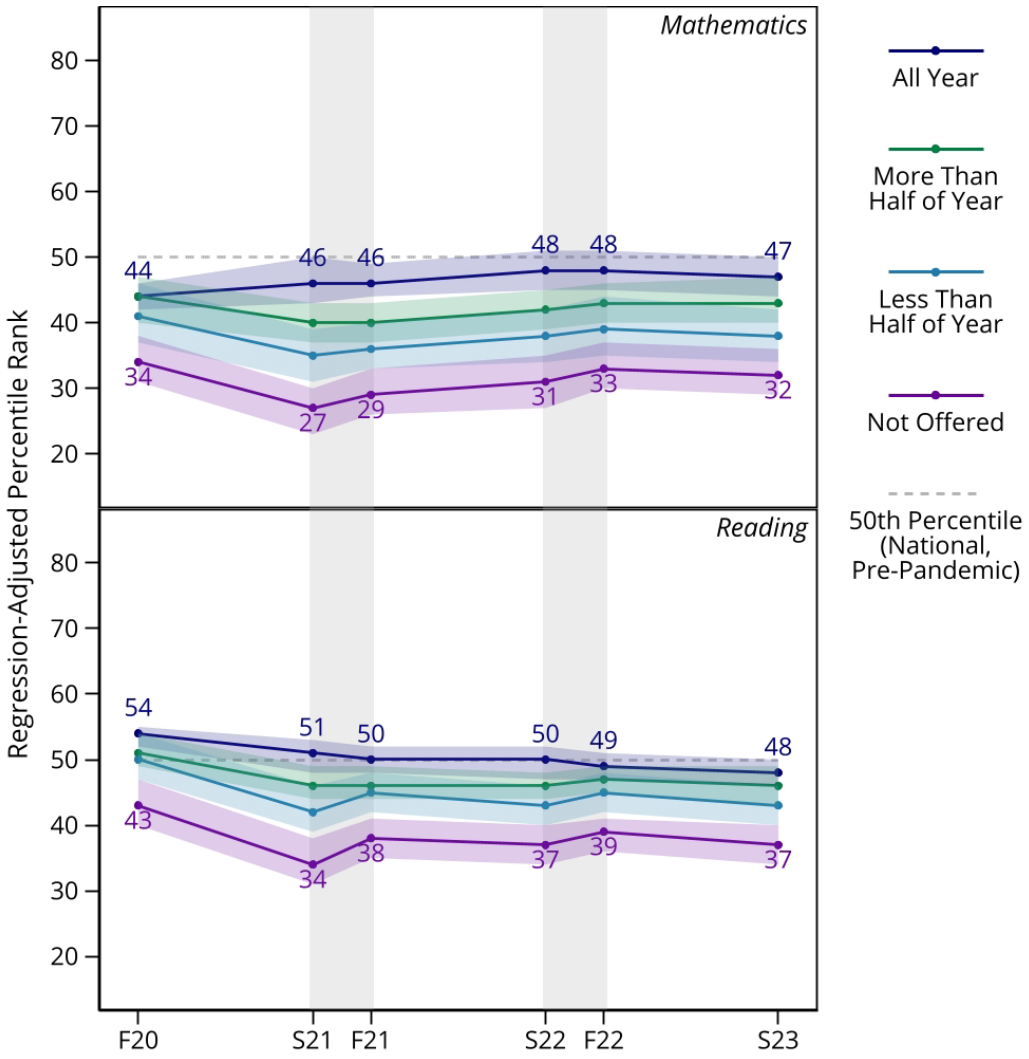
Note: These regression estimates are based on district-grade-average scores across students with MAP Growth or i-Ready scores in every possible testing period. We standardized scores relative to pre-pandemic national norms for each assessment and converted all estimates into percentile ranks. Models control for student demographics.

Figure 3.9. Regression-Adjusted Percentile Ranks by Semester and Race/Ethnicity (MAP Growth & i-Ready, 3rd-8th Grade)



Note: These regression estimates are based on district-grade-average scores across students with MAP Growth or i-Ready scores in every possible testing period. We standardized scores relative to pre-pandemic national norms for each assessment and converted all estimates into percentile ranks. Models control for student demographics. We do not show results for students who are Asian, Native American, Hawaiian or Pacific Islander, or two or more races due to low sample sizes within some of the grade levels and assessment vendors in our analysis.

Figure 3.10. Regression-Adjusted Percentile Ranks by Semester and Access to In-Person Instruction in 2020-21 (MAP Growth & i-Ready, 3rd-8th Grade)



Note: These regression estimates are based on district-grade-average scores across students with MAP Growth or i-Ready scores in every possible testing period. We standardized scores relative to pre-pandemic national norms for each assessment and converted all estimates into percentile ranks. Models control for student demographics. The modality data from 2020-21 summarize what modes of instruction districts offered during each month of the school year.

Across all testing periods, percentile ranks have consistently been the highest for districts that were in-person all year (all 9 months from September 2020 to May 2021), followed by those that were in-person for more than half (5-8 months) of the year, and those that were in-person for less than half (1-4 months) of the year, and lowest for

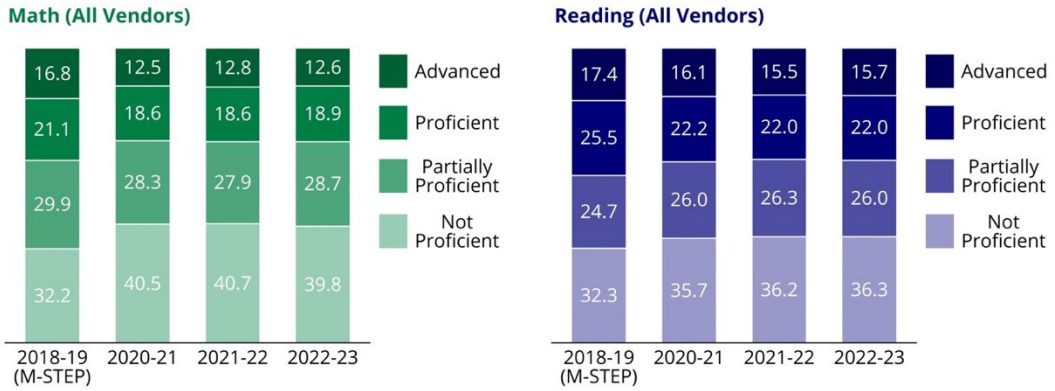
districts that did not offer in-person instruction at all in 2020-21. Gaps in reading achievement across districts that were in-person all year and those that were in-person for part of the year were smaller than the gaps between districts that were in-person for part of the year and those that were not in-person at all. Math gaps, on the other hand, are similar in size between all consecutive modality categories. In other words, math achievement gaps between districts that were in-person all 9 months and those that were in-person 5-8 months are similar in size to the gaps between districts that were in-person for 5-8 months and those that were in-person for 1-4 months, as well as the gaps between districts that were in-person for 1-4 months and those that were not in-person at all.

PROFICIENCY RATES

For a general understanding of how Michigan student performance on benchmark assessments compares to state standards for grade-level proficiency, we map student benchmark assessment scores to approximate M-STEP proficiency levels. Figure 3.11 shows the percentages of Michigan 3rd-7th grade students who are classified into each of the four M-STEP proficiency levels, based on scale score ranges that each assessment provider uses to map student benchmark assessment scores to equivalent M-STEP proficiency categories. To understand how these proficiency rates compare to similar students' performance pre-pandemic, we also show the actual proficiency rates from the 2018-19 M-STEP among all students in the districts that provided benchmark assessment data for the 2022-23 school year[7].

It is clear from Figure 3.11 that proficiency rates for Michigan students shifted substantially between the 2018-19 and 2020-21 school years. Both in the combined benchmark assessment sample and separate samples for the different benchmark assessments (shown in Appendix Figures A.1.1 to A.1.4), more students were "not proficient" and fewer were "advanced" or "proficient" in pandemic-affected years than on the 2018-19 M-STEP[8]. However, the distribution of students across the four proficiency levels changes very little after 2020-21. This suggests that, even though we see some evidence of learning recovery in later years based on trends in average scale scores and percentile ranks, these changes are not large enough in magnitude to make a substantial difference in terms of student proficiency. In other words, much larger or longer-term improvements in student achievement would be necessary to achieve the same statewide proficiency rates as in 2018-19.

Figure 3.11. M-STEP Proficiency Levels and Vendor-Defined Benchmark Assessment Equivalencies, Grades 3-7 (All Vendors)



Note: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meets the requirements in the Return to Learn law. Percentages do not always add up to 100 due to rounding.

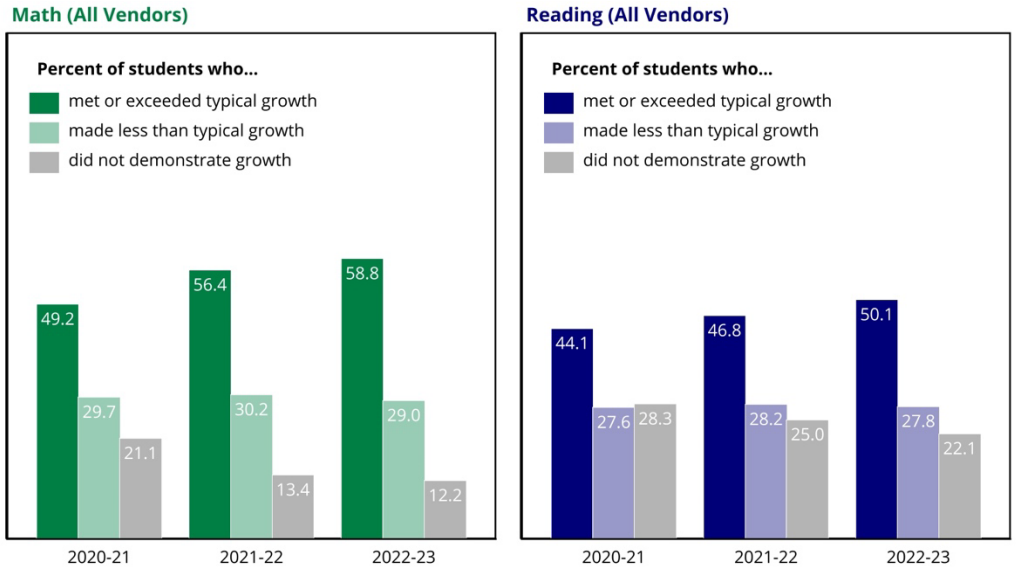
STUDENT GROWTH

Considering how much variation we observed in Michigan student achievement levels in Figures 3.4 to 3.6, changes in average achievement over time may not fully reflect the types or extents of growth that Michigan students experienced over the course of each school year. To understand the types of and variation in students' learning progress, we examine the percentage and characteristics of students who exhibited different types of growth patterns in each school year. We compare these patterns to pre-pandemic norms for "typical growth" and assess how they have changed for Michigan students from year to year.

For these analyses, we define "typical growth" as the median increase in scale scores between the fall and spring testing periods of the same school year for students from the pre-pandemic national norming samples who were in the same grade level, took the same assessment, and had similar baseline scores on their fall assessments. Thus, before the COVID-19 pandemic, we would expect to see about 50% of students meet or exceed "typical growth" each year[9]. Figures 3.12-3.17 show the percentages of students who met or exceeded typical growth (i.e., the increases to their scale scores met or surpassed the growth norm for students in their grade level with similar prior achievement scores), students who made less than typical growth (i.e., their scale scores increased by less than the pre-pandemic growth norm for students in their grade with similar prior scores), and students who did not demonstrate growth at all (i.e., their scale scores either did not change or decreased from fall to spring).

As we showed in prior reports, students in 2020-21 were less likely to meet or exceed typical growth and more likely to not demonstrate any growth at all, compared to students across the country before the COVID-19 pandemic. While more students reached their growth targets and fewer made no growth at all in 2021-22 than in 2020-21, the percentage who did not demonstrate any growth still exceeded pre-pandemic norms. Figure 3.12 shows that, in 2022-23, there were further increases in the percentage of students who met or exceeded typical growth and further decreases in the percentage who did not demonstrate growth at all.

Figure 3.12. Fall-to-Spring Growth Outcomes by School Year (All Vendors)



Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment.

In both subject areas, more than 50% of students met or exceeded their growth targets in 2022-23. This suggests that Michigan students were at least as likely as students in the pre-pandemic norming samples to achieve this level of growth. However, as we observed in the trends in average scale scores and relative percentile ranks earlier in this section, this does not mean that students have caught up to pre-pandemic achievement *levels*. Student growth targets are based on where they started at the beginning of each school year. Thus, if students are already behind grade-level standards in the fall, achieving a typical year's growth is not enough to both "catch up" to standards for students entering their current grade level *and* learn all the new

content expected for students advancing to the next grade level. For example, if a student begins the school year half a grade level behind, they would need to grow by one and one-half grade-levels to "catch up" by the end of the year.

While the share of students who did not demonstrate growth during each pandemic-affected school year exceeded pre-pandemic norms, these were generally not the same students every year. For the sample of students who have growth data from the 2020-21, 2021-22, and 2022-23 school years, Table 3.1 shows the percentage who had each possible combination of growth outcomes across the three years. Very few students—less than 1% in math and about 2% in reading—did not demonstrate growth in any of the three school years. However, only 18% met or exceeded typical math growth in all three years, and only 12% did so in reading all three years.

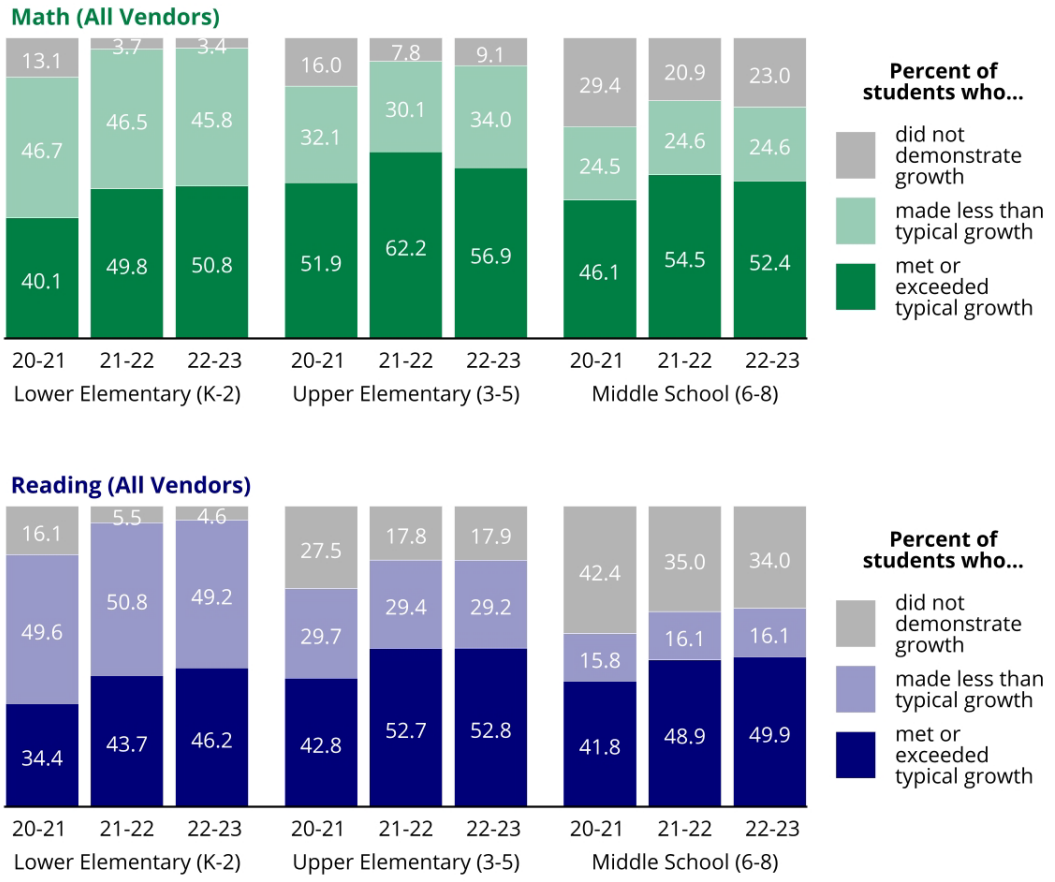
In addition to these general trends across all four MDE-approved benchmark assessments, we provide results separately for each assessment in Appendix Figures A.5.1-A.5.4. Growth outcomes for i-Ready students follow a consistent pattern as the overall trends, except that the increases in students meeting or exceeding growth targets and decreases in students not demonstrating growth are much larger than those across the combined sample in Figure 3.12. Results for the MAP Growth, Star 360, and Smarter Balanced ICA assessments, on the other hand, show similar types of changes between 2020-21 and 2021-22, but not between 2021-22 and 2022-23. This suggests that the patterns among the combined sample are driven largely by the substantial improvements in i-Ready districts. As students in i-Ready districts were less likely to have access to in-person instruction in 2020-21, their initial declines and recovery efforts were likely different from those of districts that operated in-person in 2020-21. Results for the other three assessments are consistent with national trends on the MAP Growth assessment (Lewis & Kuhfeld, 2023).

As Figure 3.13 shows, these patterns are most consistent among students in lower elementary grades. One reason for this is simply that this grade band includes new cohorts of students who did not experience the same disruptions to their learning as those who were enrolled during the school closures in 2019-20 or changes in instructional modality in 2020-21. For students in upper elementary and middle school grades, we see consistent improvements between 2020-21 and 2021-22, but little change between 2021-22 and 2022-23 (similar to the overall patterns for MAP Growth, Star 360, and Smarter Balanced ICA districts). We show additional breakdowns for each individual grade level, overall and separately by vendor for MAP Growth and i-Ready, in Appendix Figures A.6.1 to A.6.3[10].

Table 3.1. Percent of Students With Each Possible Combination of 2020-21, 2021-22, and 2022-23 Growth Outcomes				
Growth Outcome in Each School Year			Percent of Students	
2020-21	2021-22	2022-23	Math	ELA
Met or Exceeded Typical Growth	Met or Exceeded Typical Growth	Met or Exceeded Typical Growth	17.5%	12.0%
		Made Less than Typical Growth	8.8%	4.8%
		Did Not Demonstrate Growth	3.7%	4.9%
	Made Less Than Typical Growth	Met or Exceeded Typical Growth	8.0%	6.0%
		Made Less than Typical Growth	4.5%	3.3%
		Did Not Demonstrate Growth	2.0%	2.3%
	Did Not Demonstrate Growth	Met or Exceeded Typical Growth	2.1%	3.9%
		Made Less than Typical Growth	1.0%	1.5%
		Did Not Demonstrate Growth	0.9%	2.3%
Made Less Than Typical Growth	Met or Exceeded Typical Growth	Met or Exceeded Typical Growth	12.2%	9.3%
		Made Less than Typical Growth	6.8%	4.8%
		Did Not Demonstrate Growth	2.5%	3.4%
	Made Less Than Typical Growth	Met or Exceeded Typical Growth	7.0%	6.3%
		Made Less than Typical Growth	4.4%	4.4%
		Did Not Demonstrate Growth	1.5%	2.1%
	Did Not Demonstrate Growth	Met or Exceeded Typical Growth	1.5%	2.6%
		Made Less than Typical Growth	0.8%	1.2%
		Did Not Demonstrate Growth	0.6%	1.5%
Did Not Demonstrate Growth	Met or Exceeded Typical Growth	Met or Exceeded Typical Growth	4.1%	5.8%
		Made Less than Typical Growth	2.3%	2.4%
		Did Not Demonstrate Growth	1.3%	3.1%
	Made Less Than Typical Growth	Met or Exceeded Typical Growth	2.4%	2.9%
		Made Less than Typical Growth	1.5%	1.7%
		Did Not Demonstrate Growth	0.8%	1.5%
	Did Not Demonstrate Growth	Met or Exceeded Typical Growth	1.0%	2.6%
		Made Less than Typical Growth	0.4%	1.0%
		Did Not Demonstrate Growth	0.5%	1.7%

Notes: Percentages only include students with fall-to-spring growth data for all three school years. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment.

Figure 3.13. Fall-to-Spring Growth Outcomes by School Year and Grade Range (All Vendors)



Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

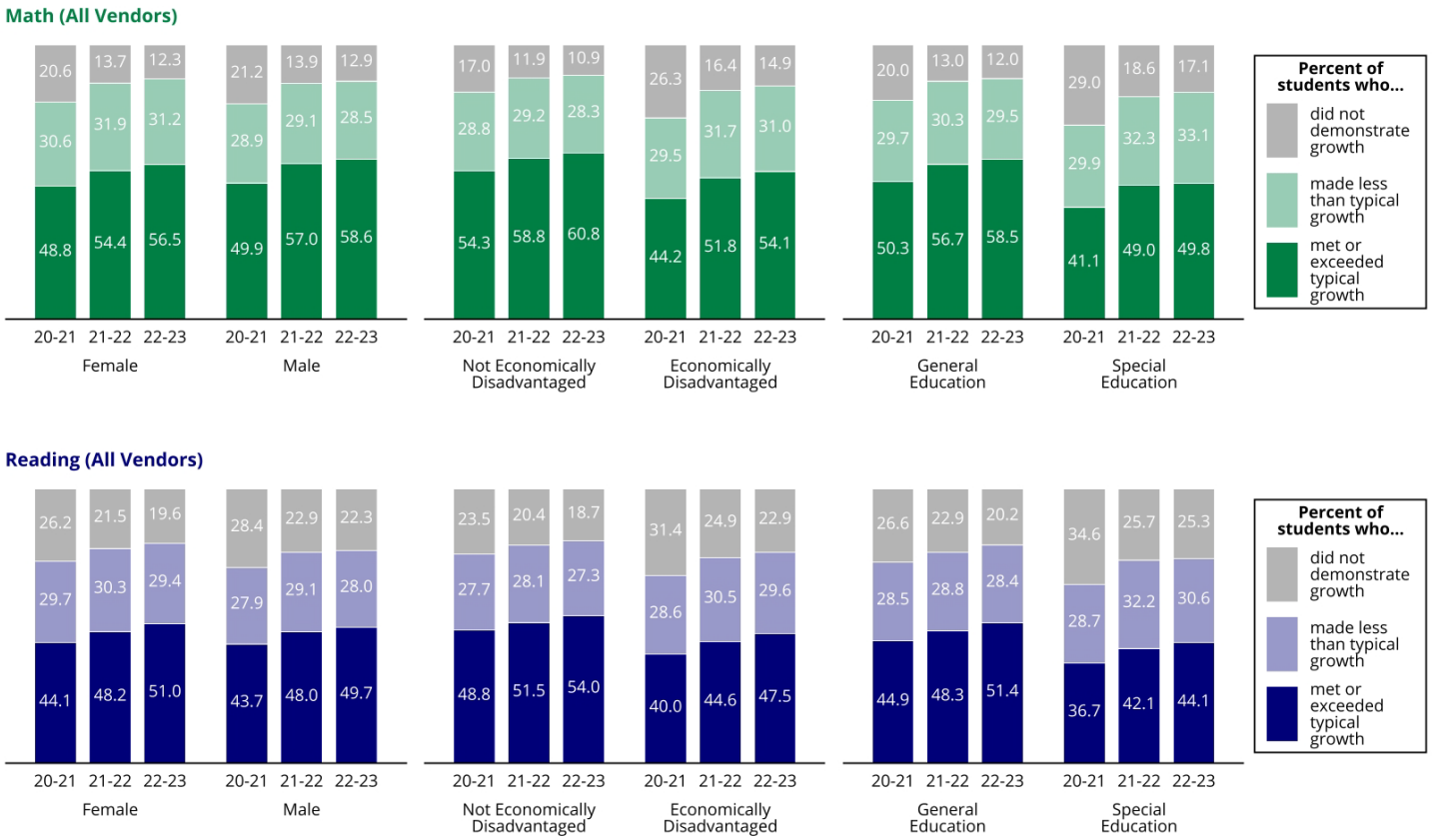
We generally find consistent patterns in student growth across demographic subgroups and other student characteristics. Across genders, economic disadvantage, and special education status, there are consistent increases from year to year in the percentage of students meeting or exceeding typical growth and decreases in the percentage not demonstrating any growth at all (shown in Figure 3.14). While the same types of patterns generally hold across races/ethnicities (shown in Figure 3.15), growth outcomes for students of color improved more than those of White students between 2020-21 and 2021-22 then remained relatively stagnant between 2021-22 and 2022-23, while changes in White students' growth outcomes were smaller in magnitude but continued into the 2022-23 school year. However, these differences may simply be

due to the differences in demographics across benchmark assessments (e.g., districts that use the i-Ready assessments, on average, serve more students of color and were more likely to be fully remote in 2020-21). When we analyze the data separately by benchmark assessment vendor (shown in Appendix Figures A.7.1-A.7.4), we find that patterns in student growth outcomes across demographic groups generally align with the overall growth patterns for the same assessment.

Figure 3.16 shows that most of the changes in student growth outcomes took place in districts that did not offer in-person instruction in 2020-21 or only did so for part of the year. The very slight changes for students in districts that offered in-person instruction all year are likely driven by students who opted to learn in a hybrid or remote format, even though their district offered an in-person option. In all other districts, we see consistent increases in students meeting or exceeding their growth targets and decreases in students not demonstrating any growth at all between 2020-21 and 2021-22. These improvements generally continue into the 2022-23 school year in reading but not in math.

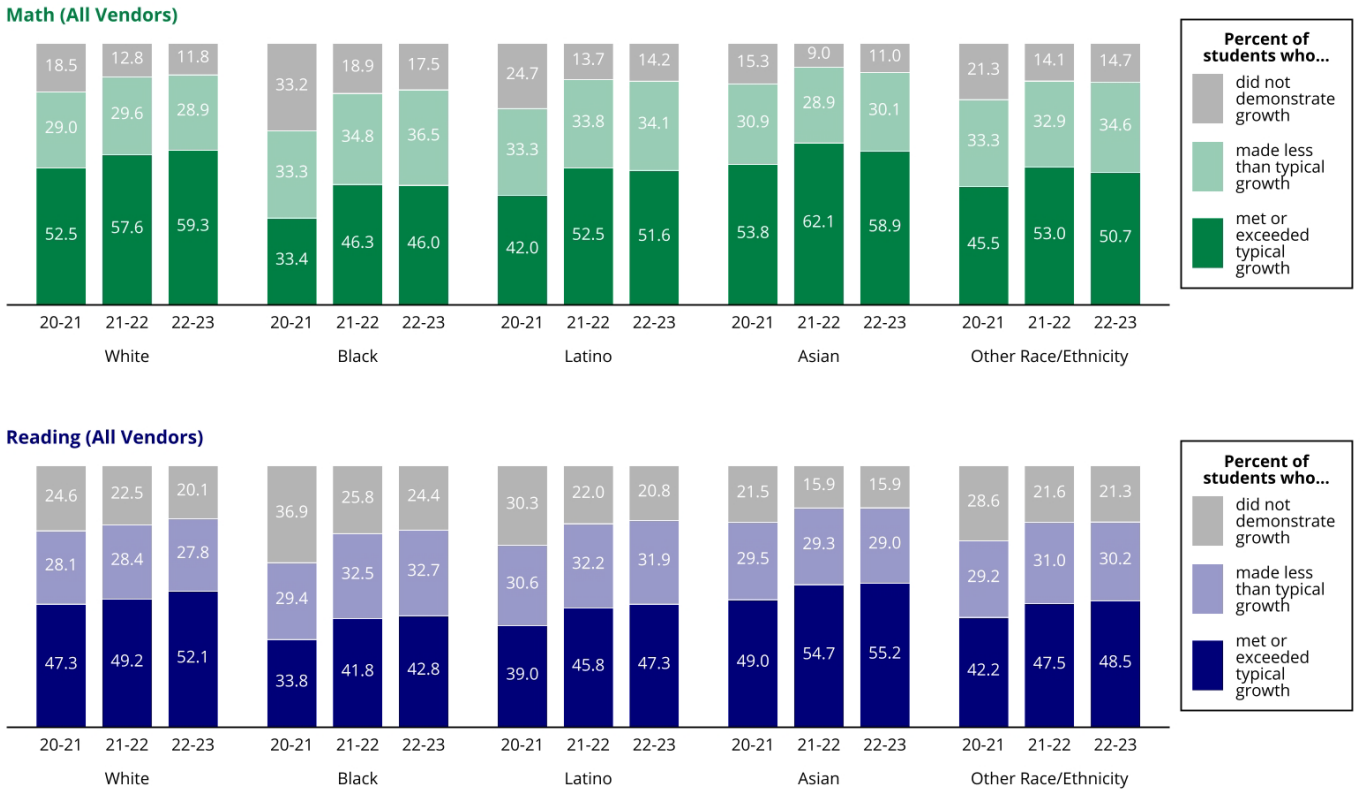
Although very few students (only about 2%) were still learning in a remote or hybrid format in 2022-23, we also compare growth outcomes for this group of students to those who learned in-person. The percentages each year are based on the subset of students who received a specific mode of instruction in 2022-23, even if that was not the same mode of instruction they received in other years. For example, the 2020-21 percentages for remote and hybrid students in Figure 3.17 show how students who received remote instruction in 2022-23 performed two years earlier. While the 98% of students who participated in in-person instruction in 2022-23 experienced improvements in their math and reading growth outcomes each year, this is not the case for the 2% who learned in a remote or hybrid format. In both math and reading, the percentages of remote and hybrid students who did not demonstrate any growth at all *increased* every year. Fewer remote and hybrid students met or exceeded typical growth in math each year, though the same was true for only reading in the first two years. These overall patterns in growth across instructional modalities generally align with those in our vendor-specific analyses (shown in Appendix Figures A.8.1-A.8.4). However, as we noted in Section Two, the majority (about 70%) of students in the "remote or hybrid" category in 2022-23 attended charter schools that have always operated virtually. Thus, these differences in growth outcomes for students who received remote or hybrid instruction in 2022-23 may simply reflect differences in the characteristics of students who enroll in virtual charter schools as opposed to differences in learning environments.

Figure 3.14. Fall-to-Spring Growth Outcomes by School Year and Student Subgroup (All Vendors)



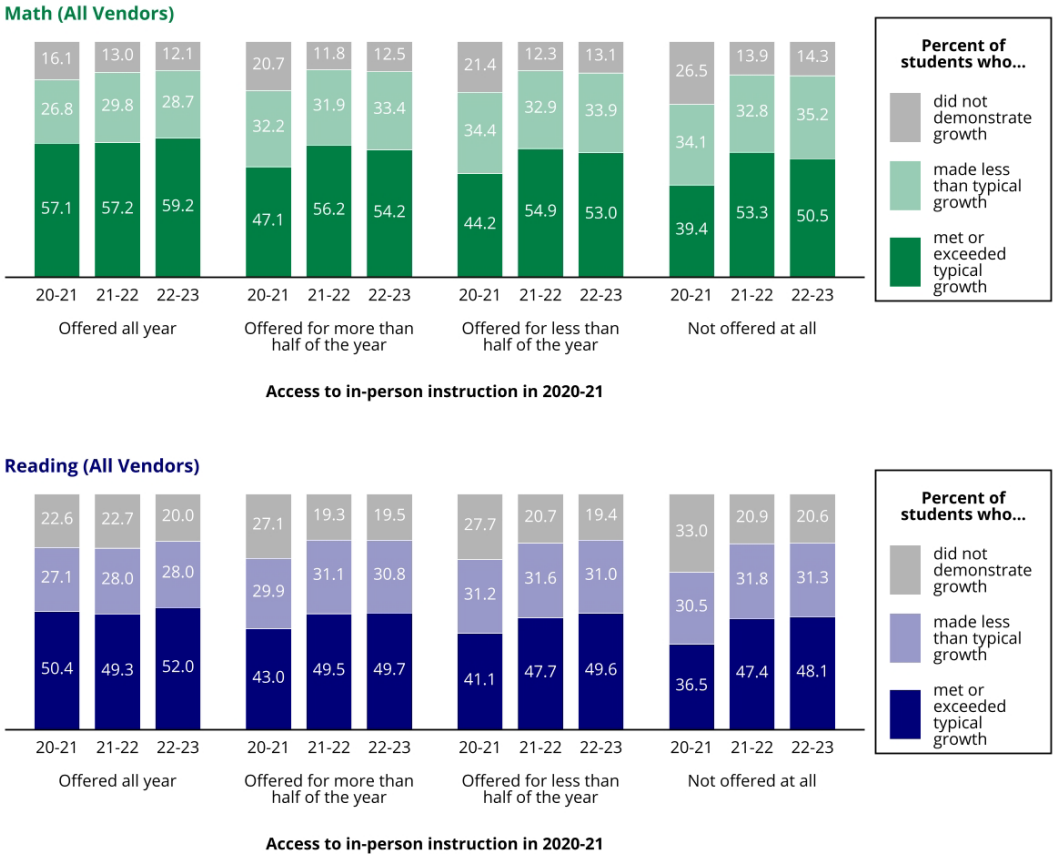
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure 3.15. Fall-to-Spring Growth Outcomes by School Year and Student Race/Ethnicity (All Vendors)



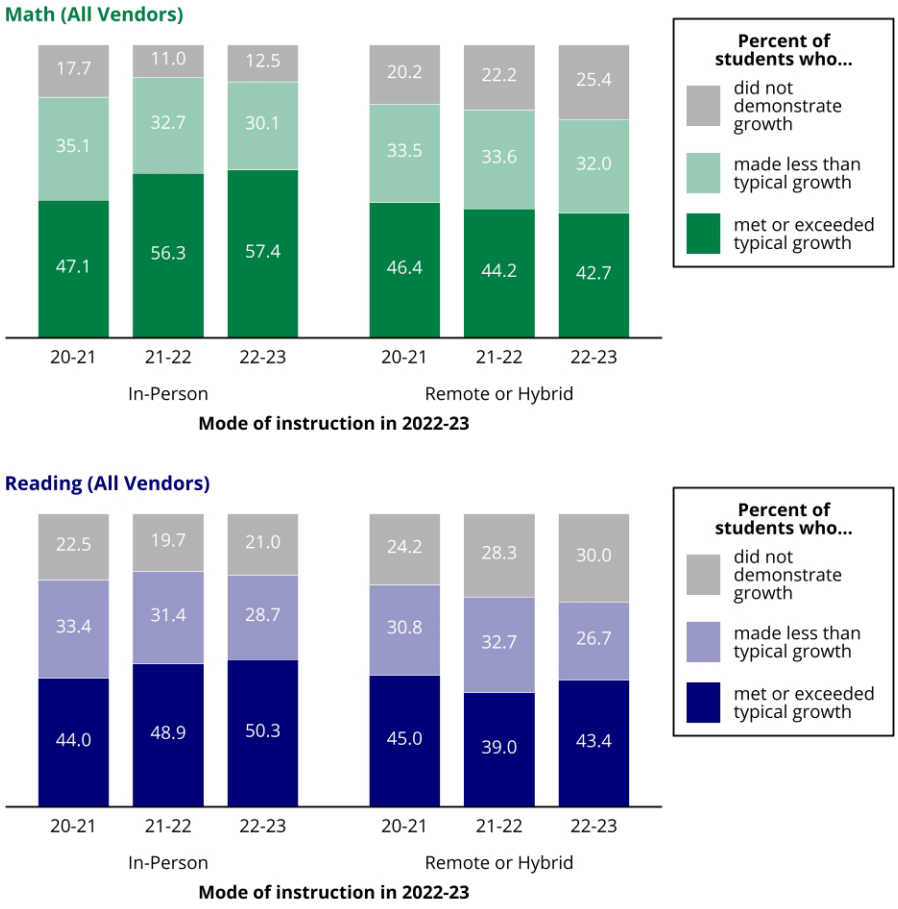
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The "Other Race/Ethnicity" category includes students who are Native American, Hawaiian or Pacific Islander, or two or more races; we cannot show separate bars for each of these groups due to low sample sizes. The percentages may not add up to 100% due to rounding.

Figure 3.16. Fall-to-Spring Growth Outcomes by School Year and Access to In-Person Instruction in 2020-21 (All Vendors)



Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The modality data from 2020-21 summarize what modes of instruction districts offered during each month of the school year. The percentages may not add up to 100% due to rounding.

Figure 3.17. Fall-to-Spring Growth Outcomes by School Year and Mode of Instruction (All Vendors)



Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The modality data from 2022-23 summarize what mode of instruction a student received during that school year. The percentages may not add up to 100% due to rounding.

Section Four: Takeaways and Implications

This report furthers our analyses of Michigan student learning during the COVID-19 pandemic by examining math and reading benchmark outcomes throughout the past three school years, from fall 2020 to spring 2023. In particular, we explore trends in average achievement and variation in student achievement compared to pre-pandemic norms, students' progress toward appropriate growth targets, and differences in academic performance across subgroups of students from different demographic groups and who experienced different modes of instruction during pandemic-affected years.

While this report helps to deepen our understanding of how Michigan public school students progressed and learned during the 2020-21, 2021-22, and 2022-23 school years, we must consider several limitations of the data when interpreting results. Most importantly, the analyses presented in this report are based on imperfect and incomplete data, as the students included in our analyses represent only a subset of the K-8 population across the state. This is notable because students who were affected the most by the pandemic may have been less likely to participate in benchmark assessments and therefore may be underrepresented in our analyses. Additionally, given that many districts administered benchmark testing virtually in the fall of 2020, it is difficult to assess fall 2020 performance and growth measures that incorporate fall 2020 achievement as a baseline to contextualize student progress. Moreover, the data available for this study does not include any prior test results for Michigan students from before fall 2020. While we can use national norms for each assessment from before the pandemic as comparison points to see how Michigan students' performance compares to students across the country in a "typical" school year, these norms may not reflect how Michigan students would have performed.

Many key findings described below show that students have started to recover academically from the school years that were most disrupted by the COVID-19 pandemic. However, there is still a long way to go to offset the tremendous effect that the COVID-19 pandemic has had on student learning. Policymakers, educators, and stakeholders should use these data to inform local and state education agencies as they continue to work to address the challenges wrought by the COVID-19 pandemic.

KEY FINDINGS

On Average, Math Achievement has Improved Slightly Since Spring 2021, While Reading Achievement Has Stayed About the Same

After accounting for differences across assessments, grade levels, and the demographic composition of students in each district, we find evidence of slight improvements in Michigan students' math achievement in 2022-23. In fall 2020, average scores for Michigan students were at about the 42nd percentile of the national norming distribution for each assessment, declining to the 39th percentile by spring 2021, and eventually returning to the 42nd percentile by spring 2023. This suggests that Michigan students' math achievement, relative to other students across the country pre-pandemic, has recovered to about the same level as it was in fall 2020. However, it is likely scores dropped between spring and fall of 2020, and hence as of spring 2023 scores probably remain below where they were before the initial school closures in 2019-20. Reading scores also declined during the 2020-21 school year, falling from the 51st to the 45th percentile of the national distribution between fall 2020 and spring 2021, but have not changed substantially since then.

Michigan Students' Achievement Levels Vary to a Greater Extent Than Would Have Been Expected Pre-Pandemic

Gaps between Michigan's highest and lowest scoring students are larger than those of students across the country who took the same assessments before the COVID-19 pandemic. In most grade levels and on most benchmark assessments, we find that the 90th percentile of Michigan students is only slightly below the 90th percentile of the national norming distribution, whereas the 10th percentile of Michigan students is much further below the 10th percentile of the national distribution. This pattern suggests that Michigan's lower-scoring students were disproportionately affected by disrupted learning experiences during the COVID-19 pandemic. These gaps between Michigan's higher and lower-scoring students have been consistent across all six testing periods between fall 2020 and spring 2023.

Students in 2022-23 Were More Likely to Reach Targets for "Typical" Growth on Their Benchmark Assessments, but Many Still did not Demonstrate Any Growth

In 2020-21, students were less likely to achieve a "typical" year's growth and more likely to not demonstrate any growth than would have been expected before the pandemic, based on national norms for each assessment. These rates have improved each year since, and in 2022-23, more than 50% of students reached or exceeded the median

growth of similar students from before the COVID-19 pandemic. However, there are still substantial percentages of students—12% in math and 22% in reading—who did not demonstrate any growth between their fall 2022 and spring 2023 benchmark assessments. These results suggest that, on average, Michigan students have returned to, and in some cases exceeded, pre-pandemic learning rates, but that many students continue to struggle.

Groups of Districts and Students That Were Most Negatively Affected by the COVID-19 Pandemic Through Spring 2021 Also Experienced the Most Learning Recovery Since Then, But Some Remain Behind

Districts that operated in a remote or hybrid format for part or all of 2020-21 were the primary drivers of the overall improvements in student growth outcomes since spring 2021. These districts are more likely to be in urban areas, serve more diverse student populations, and have more students from economically disadvantaged backgrounds than the districts that remained in person. Accelerated rates of learning in these districts led not only to improvements in overall achievement and growth outcomes at the state level, but also improvements in achievement gaps. As we showed in previous reports, some achievement gaps across races/ethnicities and students' economically disadvantaged status worsened throughout the 2020-21 school year. Following disproportionate improvements among these same subgroups of students in subsequent years, most of these achievement gaps have improved since spring 2021 and some have returned to their fall 2020 levels.

IMPLICATIONS

It Will Take More Time, Resources, and Support for Michigan Students to Recover Academically

While the evidence of improvement in student achievement and growth outcomes are encouraging, these changes are very small compared to the tremendous effect that the COVID-19 pandemic has had on student learning. To reach academic recovery, students would need to not only return to pre-pandemic learning rates but learn at accelerated rates beyond what was typical before the COVID-19 pandemic. This will not happen overnight and may not be realistic for many students before they complete their K-12 education. It will be important to not only support K-12 students in their academic recovery, but to also support students after they graduate from high school and begin their higher education.

Differentiated Instruction and Individualized Supports Will be Critical to Meet Students Where They Are

Students' academic levels and learning needs are more varied now than before the COVID-19 pandemic. School districts should ensure that teachers have access to professional development resources, curricula, and instructional materials that are appropriate for their students' current academic levels, which may be different and more varied than they were in prior years. Continually assessing and accommodating students' diverse needs and providing instructional interventions in one-on-one and small group settings are only feasible when schools can hire and maintain the necessary instructional and support staff. Michigan has recently made budget investments to enable these efforts for districts and schools, but we will need long-term commitments to this level of support, rather than relying on short-term infusions of resources, to make necessary progress.

These Unprecedented Challenges Are Widespread Both Within and Outside of Michigan

These findings are not unique to Michigan. Proficiency data from 25 different states' 2023 summative assessments show that most states are making progress in math but have not yet recovered (*State Test Score Data Briefs*, 2023). Reading recovery has been much more varied. Some states, like Michigan, experienced minimal change in reading achievement, while others made progress or even recovered, and a few experienced continued declines. As school districts and states continue their efforts to accelerate student learning, it will be important to monitor individual students' progress, identify promising practices throughout the state and country, and adjust recovery strategies as new evidence emerges about what is and isn't working for students and schools in similar contexts.

References

Bailey, D. H., Duncan, G. J., Murnane, R. J., & Au Yeung, N. (2021). Achievement gaps in the wake of COVID-19. *Educational Researcher*, *50*(5), 266–275. https://doi.org/10.3102/0013189X211011237

Barnum, M., & Belsha, K. (2023, October 2). *Blizzard of state test scores shows some progress in math, divergence in reading*. Chalkbeat. https://www.chalkbeat.org/2023/10/2/23896045/state-test-scores-data-math-reading-pandemic-era-learning-loss

Curriculum Associates. (2020a). *I-Ready diagnostic national norms tables for reading and mathematics for grades K-8 2020-21* (Research Report RR 2020-40). Curriculum Associates, LLC. https://cdn.bfldr.com/LS6J0F7/as/ck29g9k45bstghwgvs8m4xmb/iready-norms-tables-K-8

Curriculum Associates. (2020b). *Scores on i-Ready Diagnostic that are equivalent to performance levels on the Michigan Student Test of Educational Progress (M-STEP)* (Research Report RR 2020-29). Curriculum Associates, LLC.

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). COVID-19 and learning loss—Disparities grow and students need help. *McKinsey & Company, December*, *8*, 6–7.

Goldhaber, D., Kane, T. J., McEachin, A., & Morton, E. (2022). A comprehensive picture of achievement across the COVID-19 pandemic years: Examining variation in

test levels and growth across districts, schools, grades, and students. In *National Center for Analysis of Longitudinal Data in Education Research (CALDER)* (Working Paper 266–0522). National Center for Analysis of Longitudinal Data in Education Research. https://eric.ed.gov/?id=ED620384

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* (Working Paper 30010). National Bureau of Economic Research. https://doi.org/10.3386/w30010

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/mis-2020-21-and-2021-22-benchmark-assessments/

Lewis, K., & Kuhfeld, M. (2023). *Education's long COVID: 2022–23 achievement data reveal stalled progress toward pandemic recovery* [NWEA Research Brief]. Northwest Evaluation Association. https://www.nwea.org/uploads/Educations-long-covid-2022-23-achievement-data-reveal-stalled-progress-toward-pandemic-recovery_NWEA_Research-brief.pdf

McEachin, A., & Atteberry, A. (2017). The impact of summer learning loss on measures of school performance. *Education Finance and Policy*, *12*(4), 468–491. https://doi.org/10.1162/edfp_a_00213

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

NWEA. (2020). *Linking study report: Predicting performance on the Michigan state assessment system in grades 3-8 ELA and mathematics based on NWEA MAP Growth scores*. Northwest Evaluation Association. https://www.nwea.org/content/uploads/2016/12/MI-MAP-Growth-LinkingStudy-Report_NWEA_2020-12-22.pdf

Pier, L., Christian, M., Tymeson, H., & Meyer, R. H. (2021). *COVID-19 and the educational equity crisis* (pp. 1–49). Stanford Graduate School of Education. https://edpolicyinca.org/publications/covid-19-impacts-student-learning

Quinn, D. M., Cooc, N., McIntyre, J., & Gomez, C. J. (2016). Seasonal dynamics of academic achievement inequality by socioeconomic status and race/ethnicity: Updating and extending past research with new national data. *Educational Researcher*, *45*(8), 443–453. https://doi.org/10.3102/0013189X16677965

Renaissance Learning. (2019). *Pathway to proficiency: Linking Star Reading and Star Math to the Michigan Student Test of Educational Progress (M-STEP)* [Technical Paper]. Renaissance Learning, Inc. https://doc.renlearn.com/KMNet/R0053649AF33BE0B.pdf

Renaissance Star Assessments. (2021a). *Star assessments for early literacy technical manual* [Technical Manual]. Renaissance Learning, Inc. https://star-help.renaissance.com/hc/en-us/articles/12483321397019-Star-Assessments-for-Early-Literacy-Technical-Manual

Renaissance Star Assessments. (2021b). *Star assessments for math technical manual* [Technical Manual]. Renaissance Learning, Inc. https://star-help.renaissance.com/hc/en-us/articles/11583043402267-Star-Assessments-for-Math-Technical-Manual

Renaissance Star Assessments. (2021c). *Star assessments for reading technical manual* [Technical Manual]. Renaissance Learning, Inc. https://star-help.renaissance.com/hc/en-us/articles/12542471051803-Star-Assessments-for-Reading-Technical-Manual

Smarter Balanced Assessment Consortium. (2020a). *2018-19 summative technical report* [Technical Report]. University of California. https://technicalreports.smarterbalanced.org/2018-19_summative-report/_book/

Smarter Balanced Assessment Consortium. (2021). *2020-21 interim technical report* [Technical Report]. University of California. https://technicalreports.smarterbalanced.org/2020-21_interim-report/_book/

Smarter Balanced Assessment Consortium. (2020b). *Smarter Balanced validity research percentiles* [Research Report]. Smarter Balanced Validity Research. https://validity.smarterbalanced.org/reports-and-specifications/percentiles/

State test score data briefs. (2023). [dataset]. COVID-19 School Data Hub. https://www.covidschooldatahub.com/score-results

Strunk, K. O., Hopkins, B., Kilbride, T., Imberman, S., & Yu, D. (2023). *The path of student learning delay during the COVID-19 pandemic: Evidence from Michigan*. Education Policy Innovation Collaborative. https://epicedpolicy.org/wp-content/uploads/2023/05/Benchmark_PolBrief_StudentOutcomes_May2023.pdf

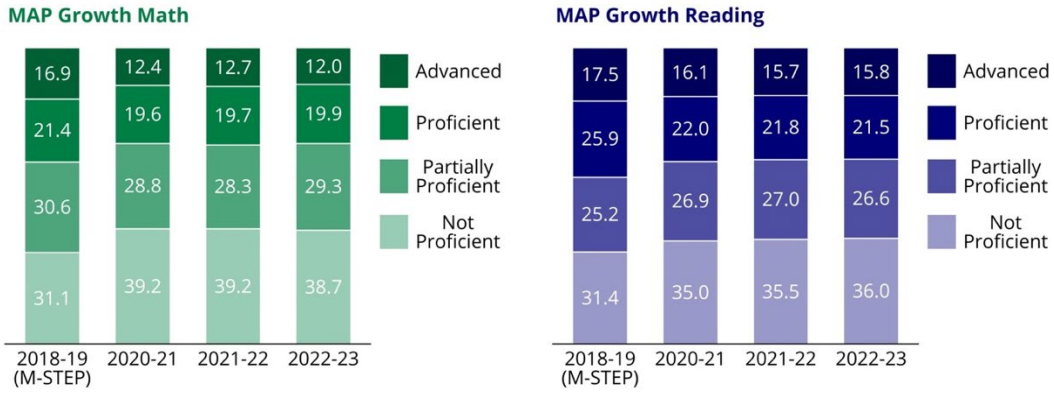
Thum, Y. M., & Kuhfeld, M. (2020). *NWEA 2020 MAP growth: Achievement status and growth norms for students and schools* (pp. 1–574) [White paper]. Northwest

Evaluation Association. https://teach.mapnwea.org/impl/normsResearchStudy.pdf

von Hippel, P. T., & Hamrock, C. (2019). Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them. *Sociological Science*, *6*(3), 43–80. https://doi.org/10.15195/v6.a3

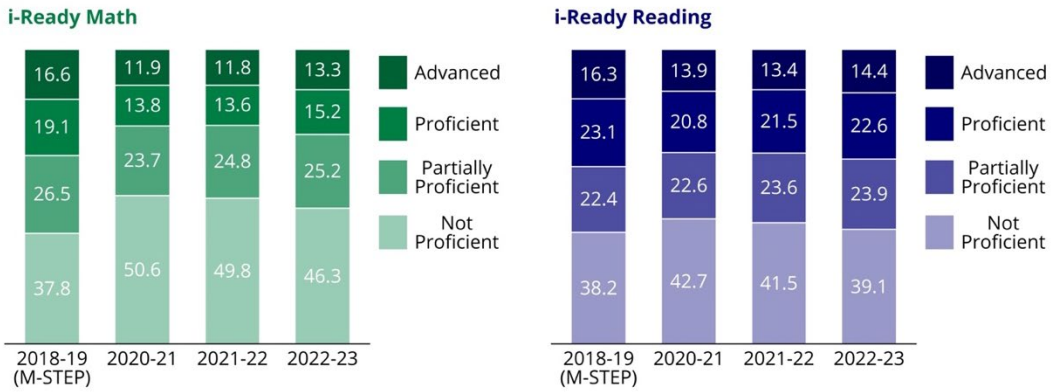
Appendix

Figure A.1.1. M-STEP Proficiency Levels and Vendor-Defined Benchmark Assessment Equivalencies, Grades 3-7 (MAP Growth)



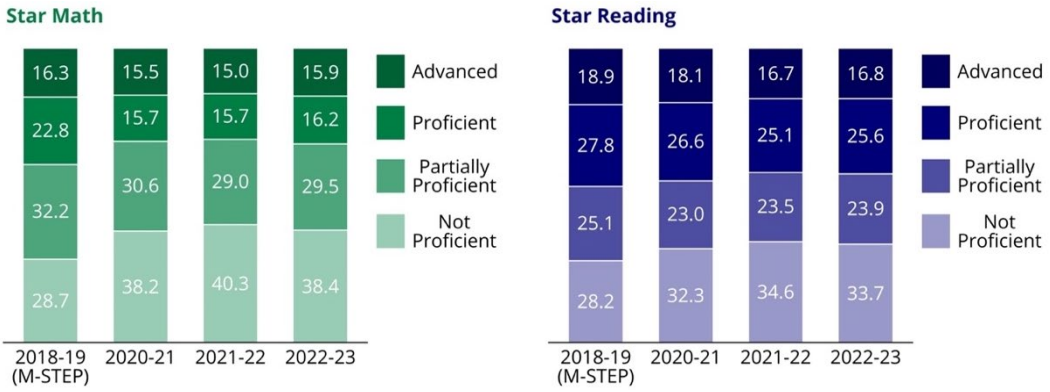
Notes: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meet the requirements in the Return to Learn law. The percentages may not add up to 100% due to rounding.

Figure A.1.2. M-STEP Proficiency Levels and Vendor-Defined Benchmark Assessment Equivalencies, Grades 3-7 (i-Ready)



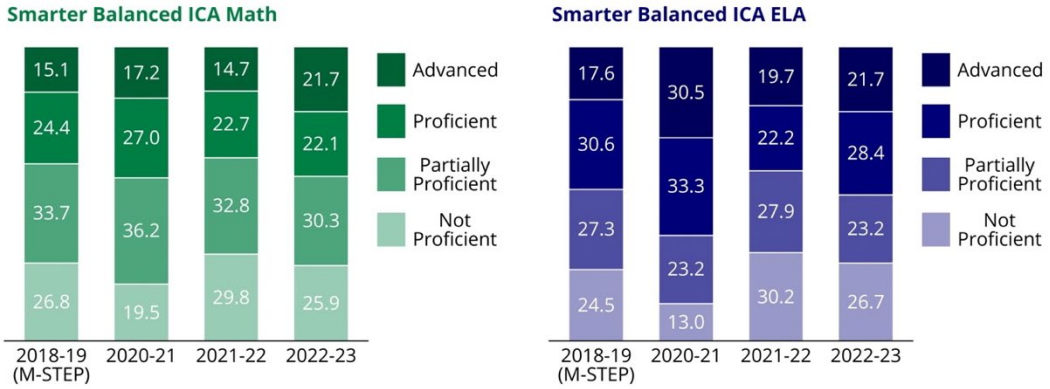
Notes: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meet the requirements in the Return to Learn law. The percentages may not add up to 100% due to rounding.

Figure A.1.3. M-STEP Proficiency Levels and Vendor-Defined Benchmark Assessment Equivalencies, Grades 3-7 (Star 360)



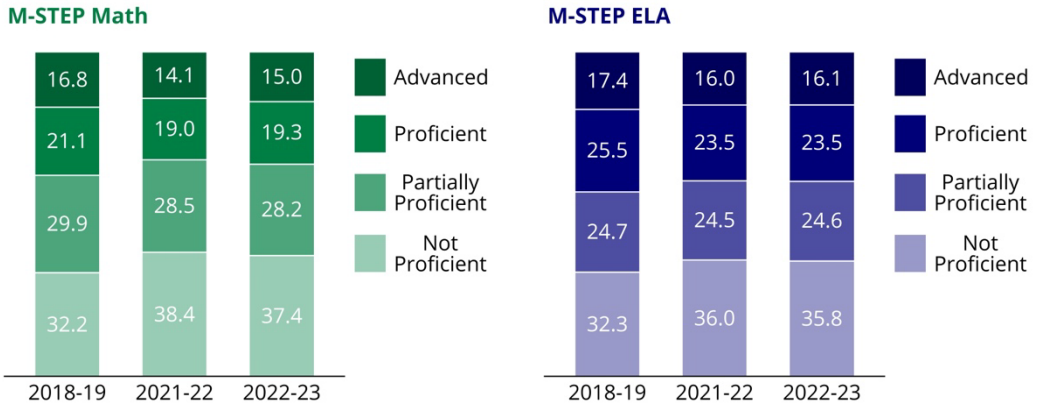
Notes: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meet the requirements in the Return to Learn law. The percentages may not add up to 100% due to rounding.

Figure A.1.4. M-STEP Proficiency Levels and Vendor-Defined Benchmark Assessment Equivalencies, Grades 3-7 (Smarter Balanced ICA)



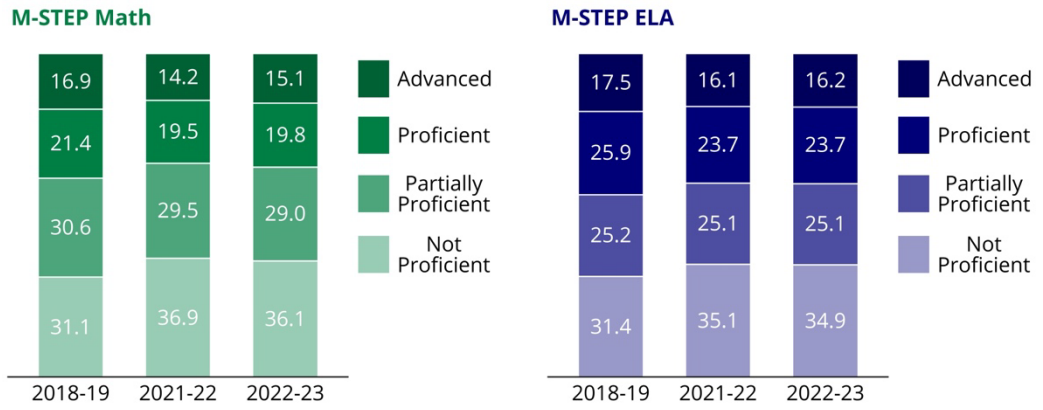
Notes: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meet the requirements in the Return to Learn law. The percentages may not add up to 100% due to rounding.

Figure A.1.5. M-STEP Proficiency Levels, Grades 3-7 (All Districts with 2022-23 Benchmark Assessment Data)



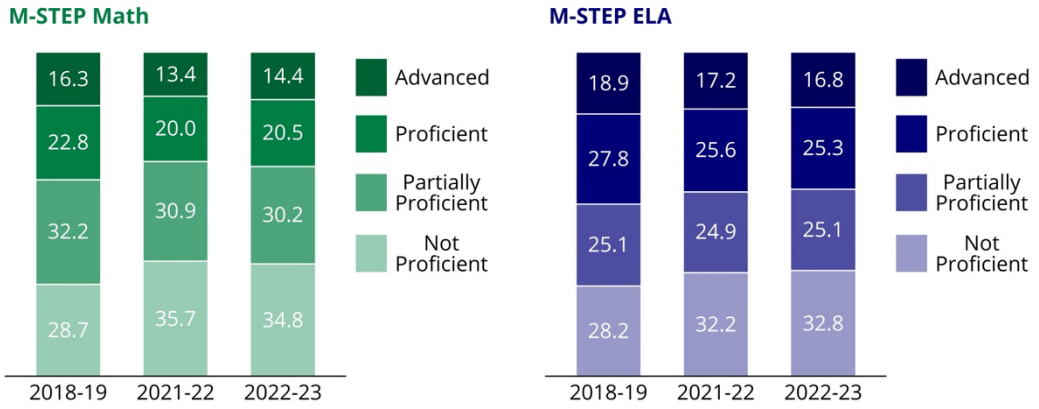
Notes: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meet the requirements in the Return to Learn law. The percentages may not add up to 100% due to rounding.

Figure A.1.6. M-STEP Proficiency Levels, Grades 3-7 (MAP Growth Districts)



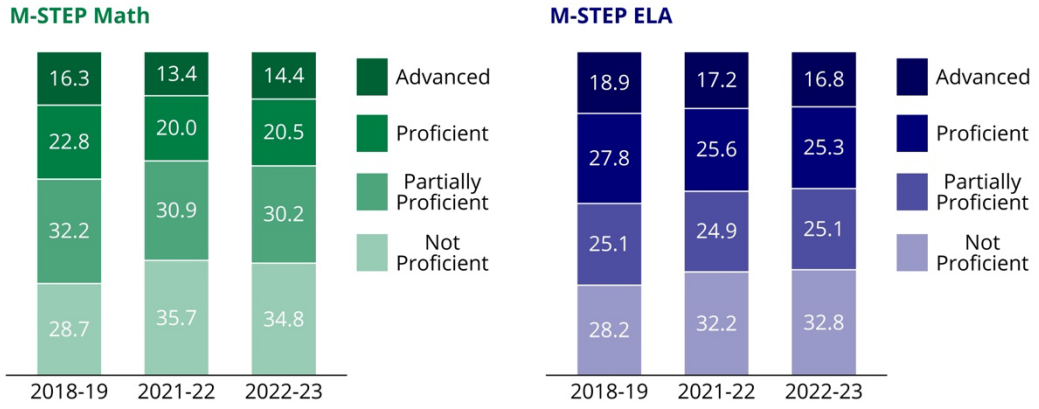
Notes: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meet the requirements in the Return to Learn law. The percentages may not add up to 100% due to rounding.

Figure A.1.7. M-STEP Proficiency Levels, Grades 3-7 (i-Ready Districts)



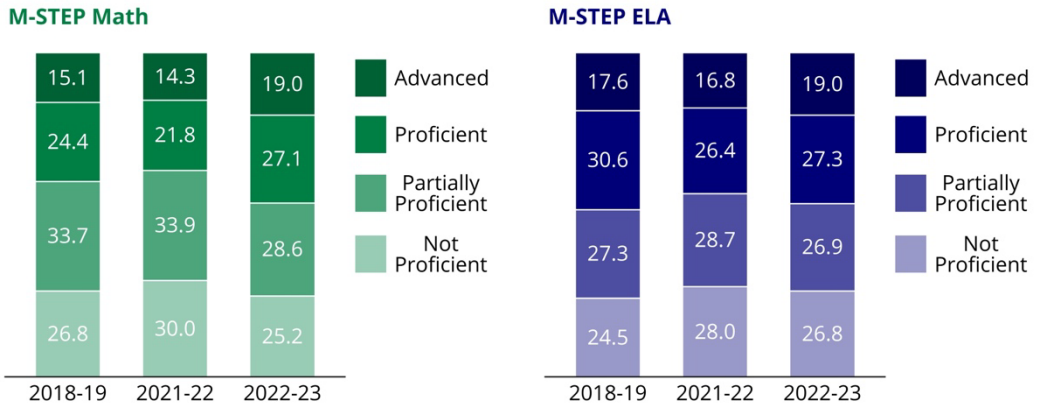
Notes: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meet the requirements in the Return to Learn law. The percentages may not add up to 100% due to rounding.

Figure A.1.8. M-STEP Proficiency Levels, Grades 3-7 (Star 360 Districts)



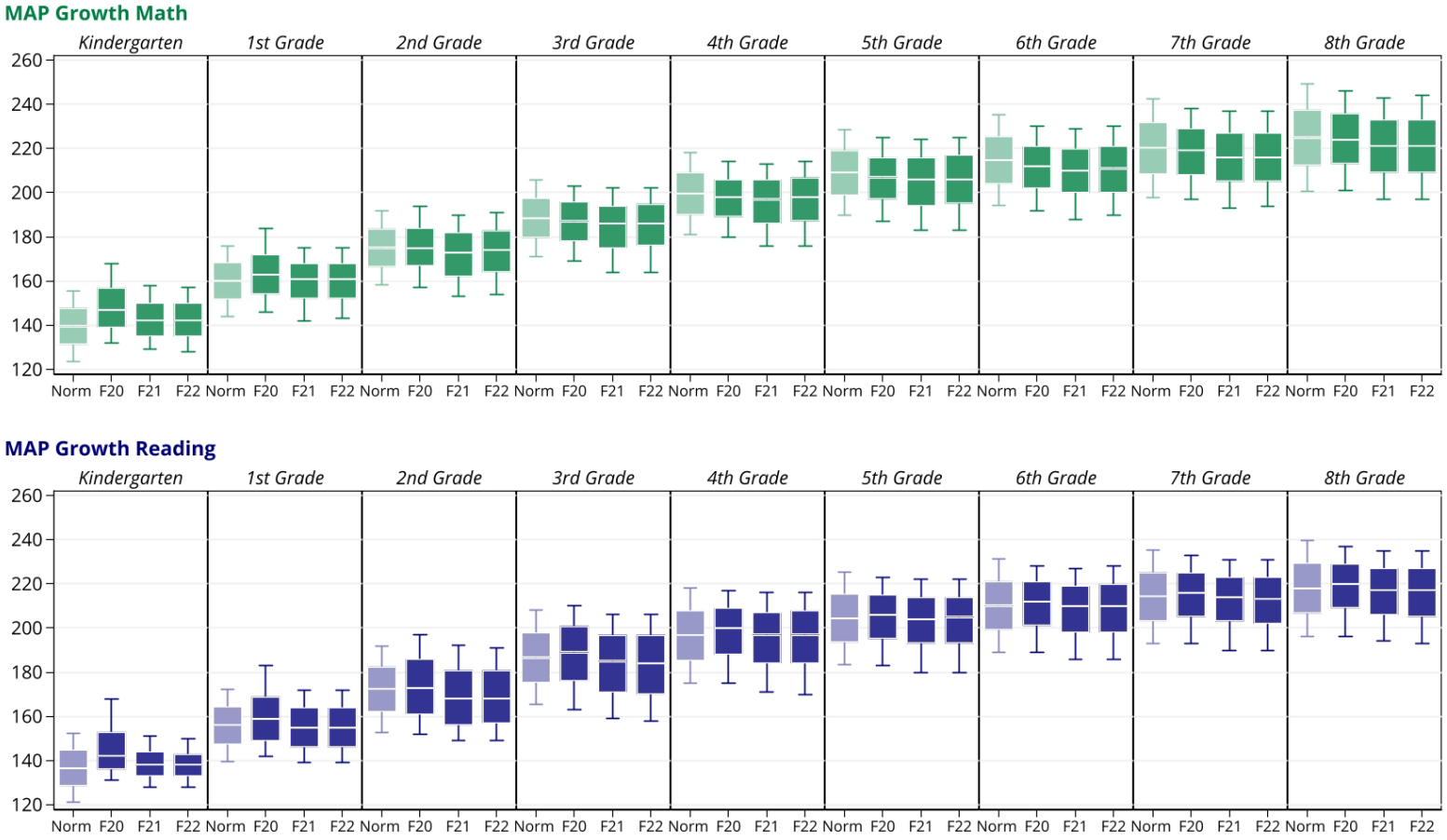
Notes: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meet the requirements in the Return to Learn law. The percentages may not add up to 100% due to rounding.

Figure A.1.9. M-STEP Proficiency Levels, Grades 3-7 (Smarter Balanced ICA Districts)



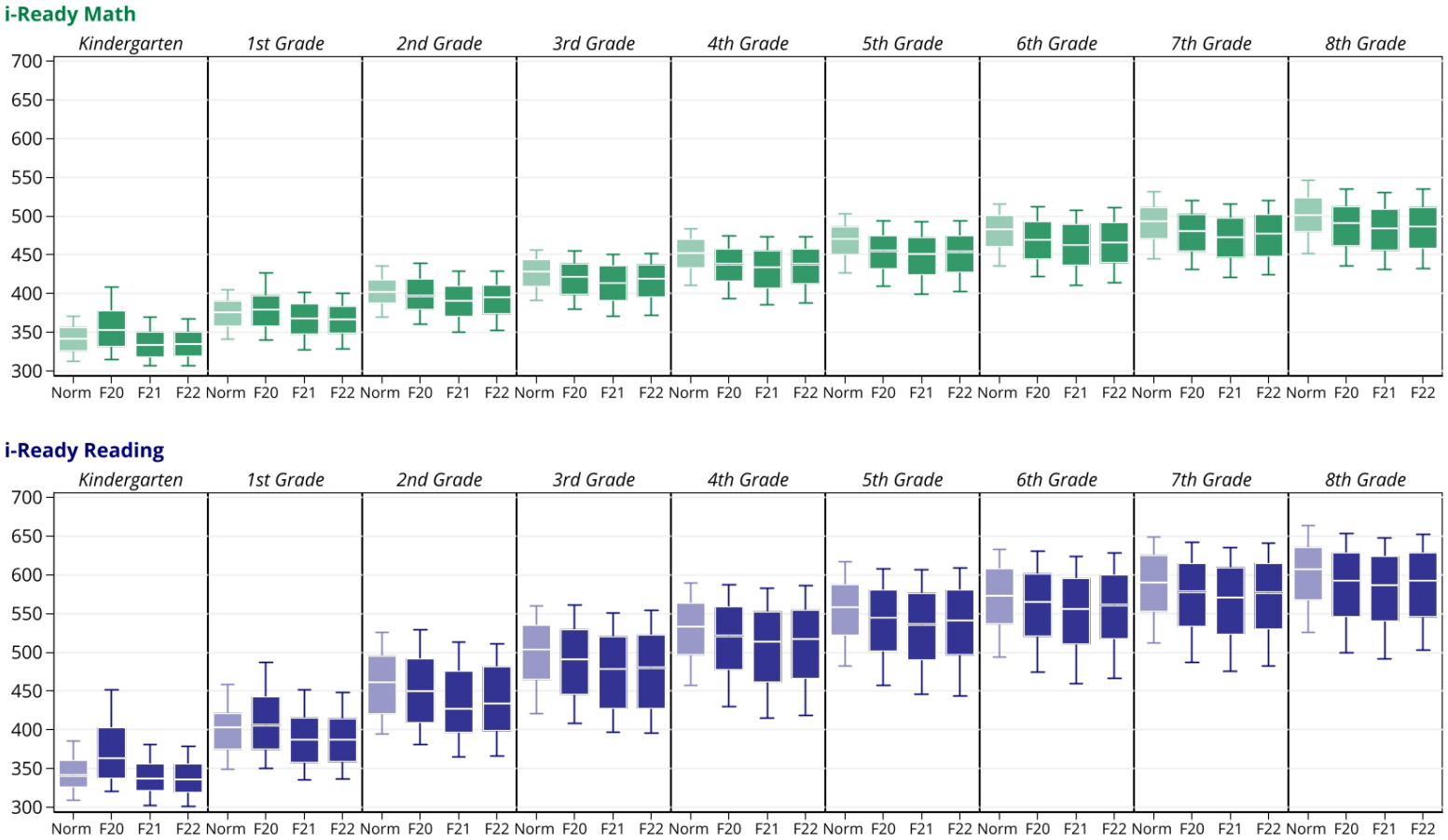
Notes: These percentages include all 3rd- through 7th-grade students in districts that provided benchmark assessment data for the 2022-23 school year that meet the requirements in the Return to Learn law. The percentages may not add up to 100% due to rounding.

Figure A.2.1. Distribution of Fall Scale Scores by School Year and Grade Level, MAP Growth



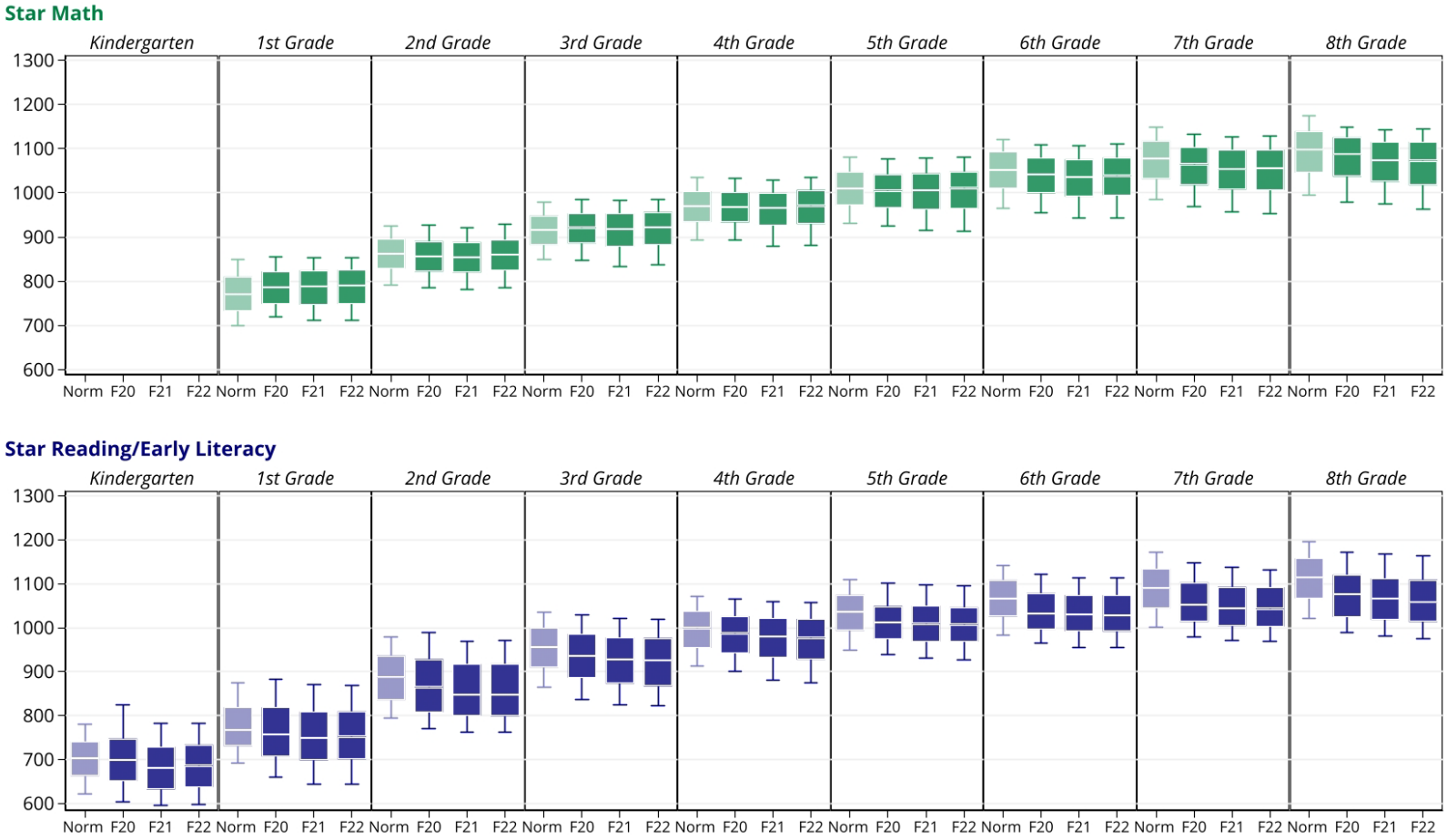
Notes: Each vertical line shows the 10th and 90th percentiles of student fall scores, while each rectangle shows the 25th, 50th, and 75th percentiles. "Norm," "F20," "F21," and "F22" represent the pre-pandemic national norm and the fall 2020, fall 2021, and fall 2022 testing periods, respectively.

Figure A.2.2. Distribution of Fall Scale Scores by School Year and Grade Level, i-Ready



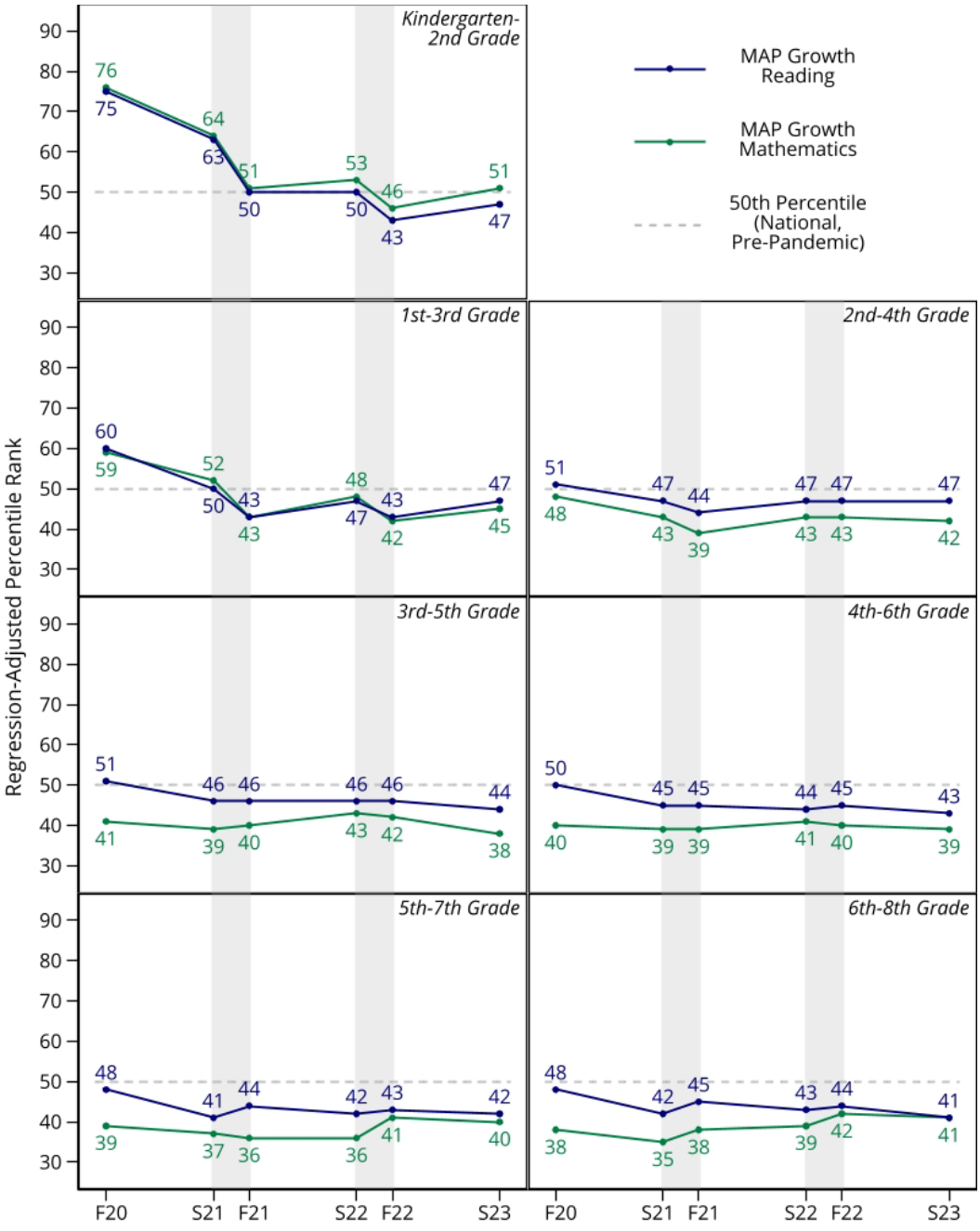
Notes: Each vertical line shows the 10th and 90th percentiles of student fall scores, while each rectangle shows the 25th, 50th, and 75th percentiles. "Norm," "F20," "F21," and "F22" represent the pre-pandemic national norm and the fall 2020, fall 2021, and fall 2022 testing periods, respectively.

Figure A.2.3. Distribution of Fall Scale Scores by School Year and Grade Level, Star 360



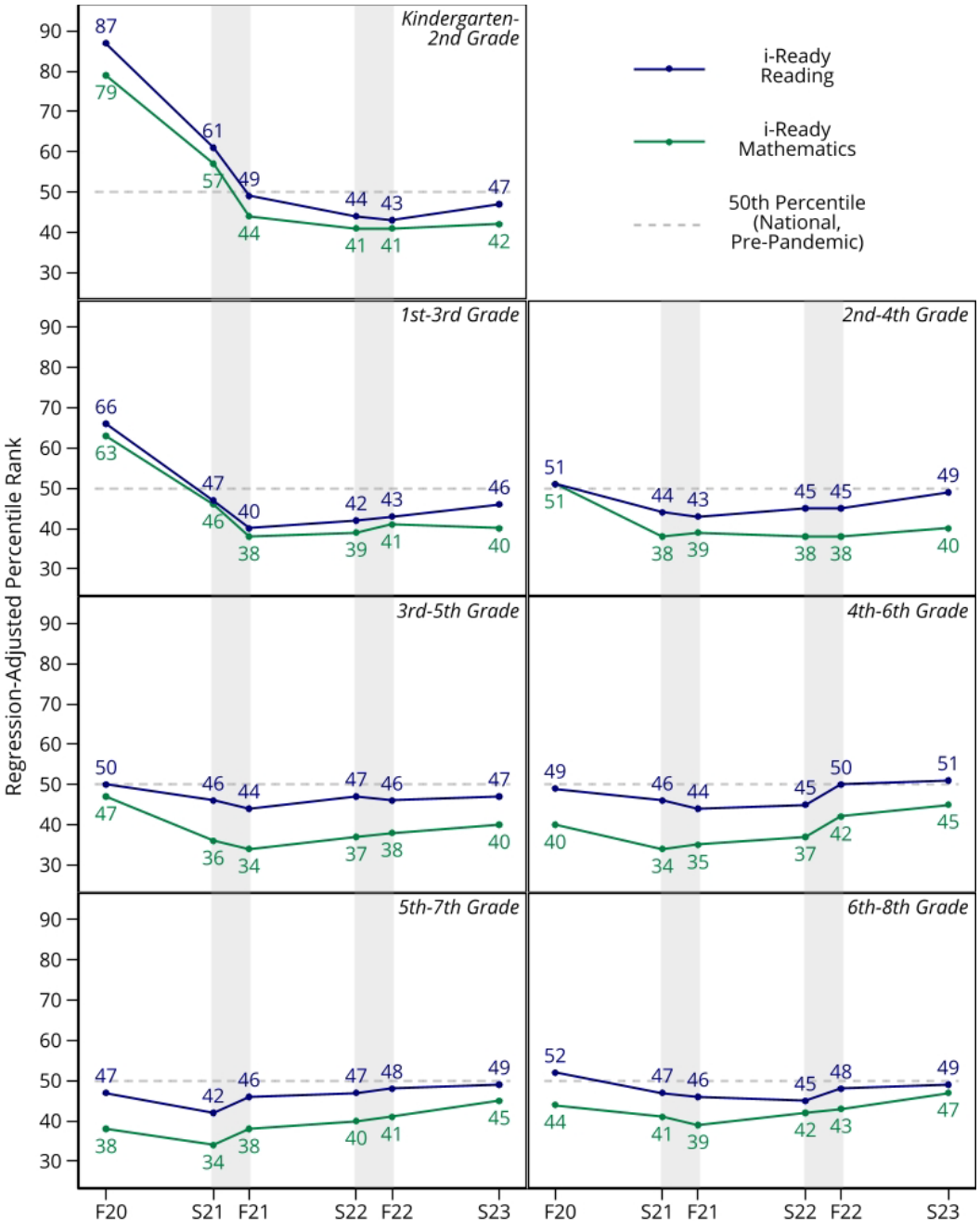
Notes: Each vertical line shows the 10th and 90th percentiles of student fall scores, while each rectangle shows the 25th, 50th, and 75th percentiles. "Norm," "F20," "F21," and "F22" represent the pre-pandemic national norm and the fall 2020, fall 2021, and fall 2022 testing periods, respectively.

Figure A.3.1 Regression-Adjusted Percentile Ranks by Semester and Grade Level (MAP Growth)



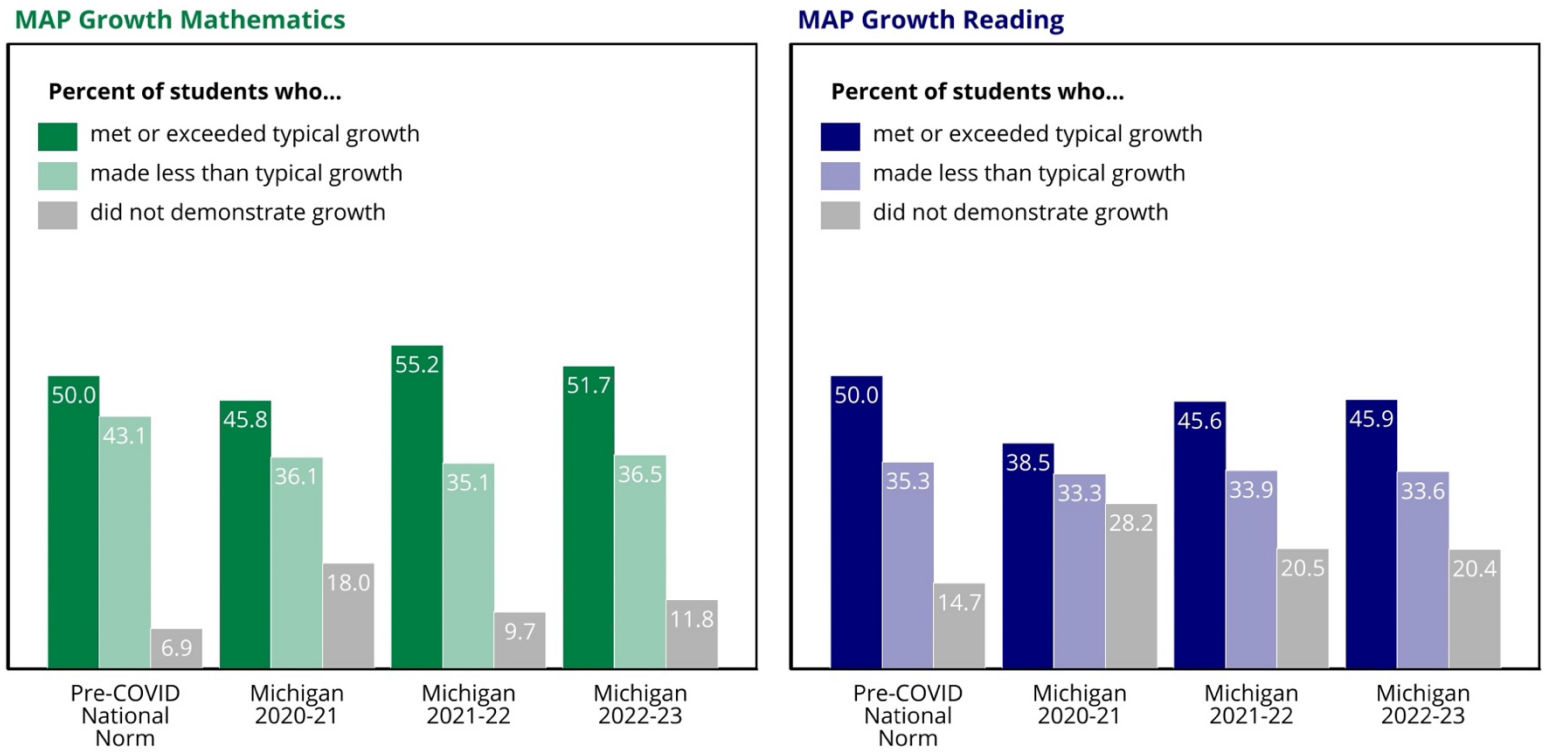
Note: These regression estimates are based on district-grade-average scores across students with MAP Growth scores in every possible testing period. We standardized scores relative to pre-pandemic national norms for each assessment and converted all estimates into percentile ranks.

Figure A.3.2. Regression-Adjusted Percentile Ranks by Semester and Grade Level (i-Ready)



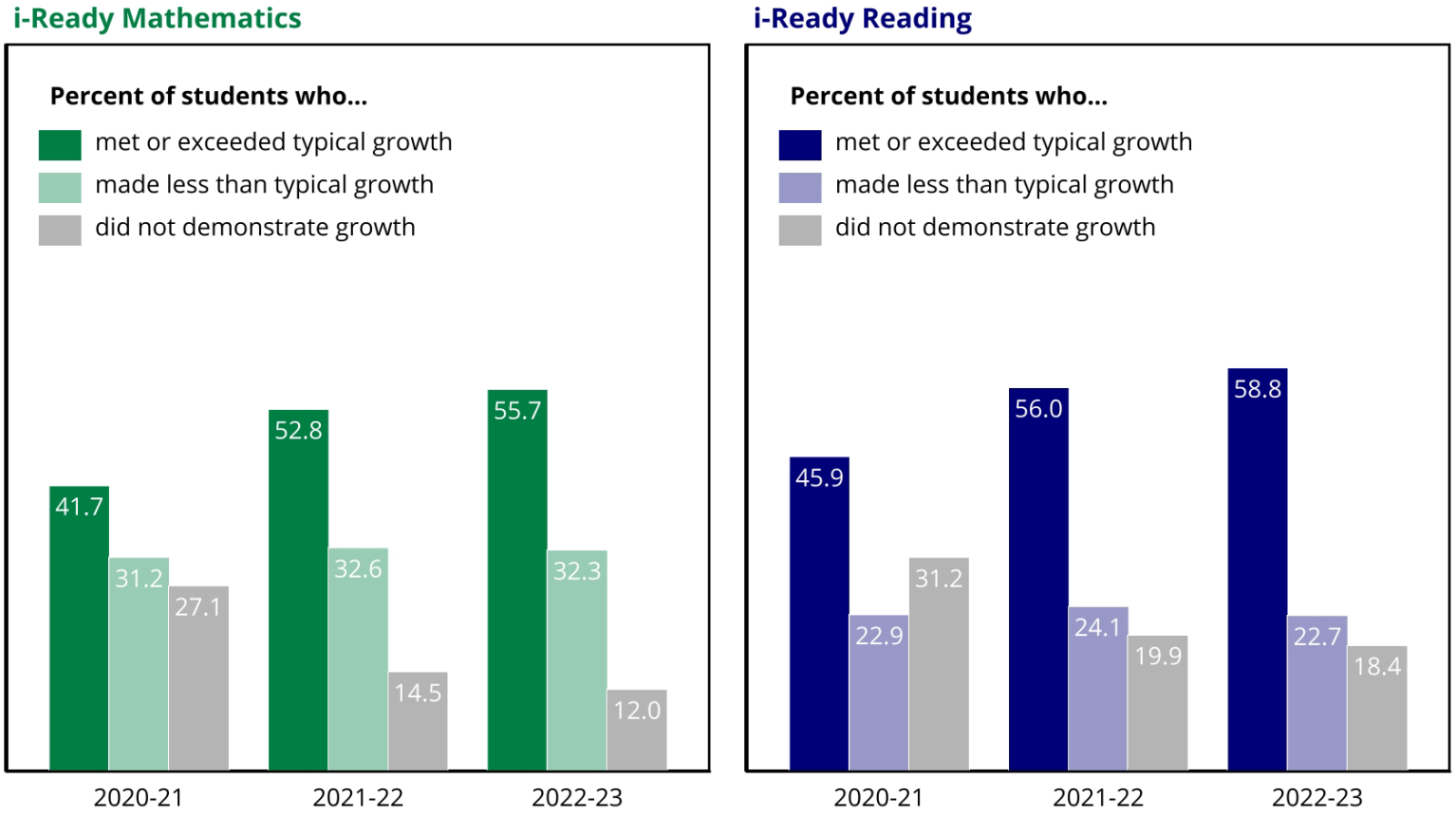
Note: These regression estimates are based on district-grade-average scores across students with i-Ready scores in every possible testing period. We standardized scores relative to pre-pandemic national norms for each assessment and converted all estimates into percentile ranks.

Figure A.5.1. Fall-to-Spring Growth Outcomes by School Year (MAP Growth)



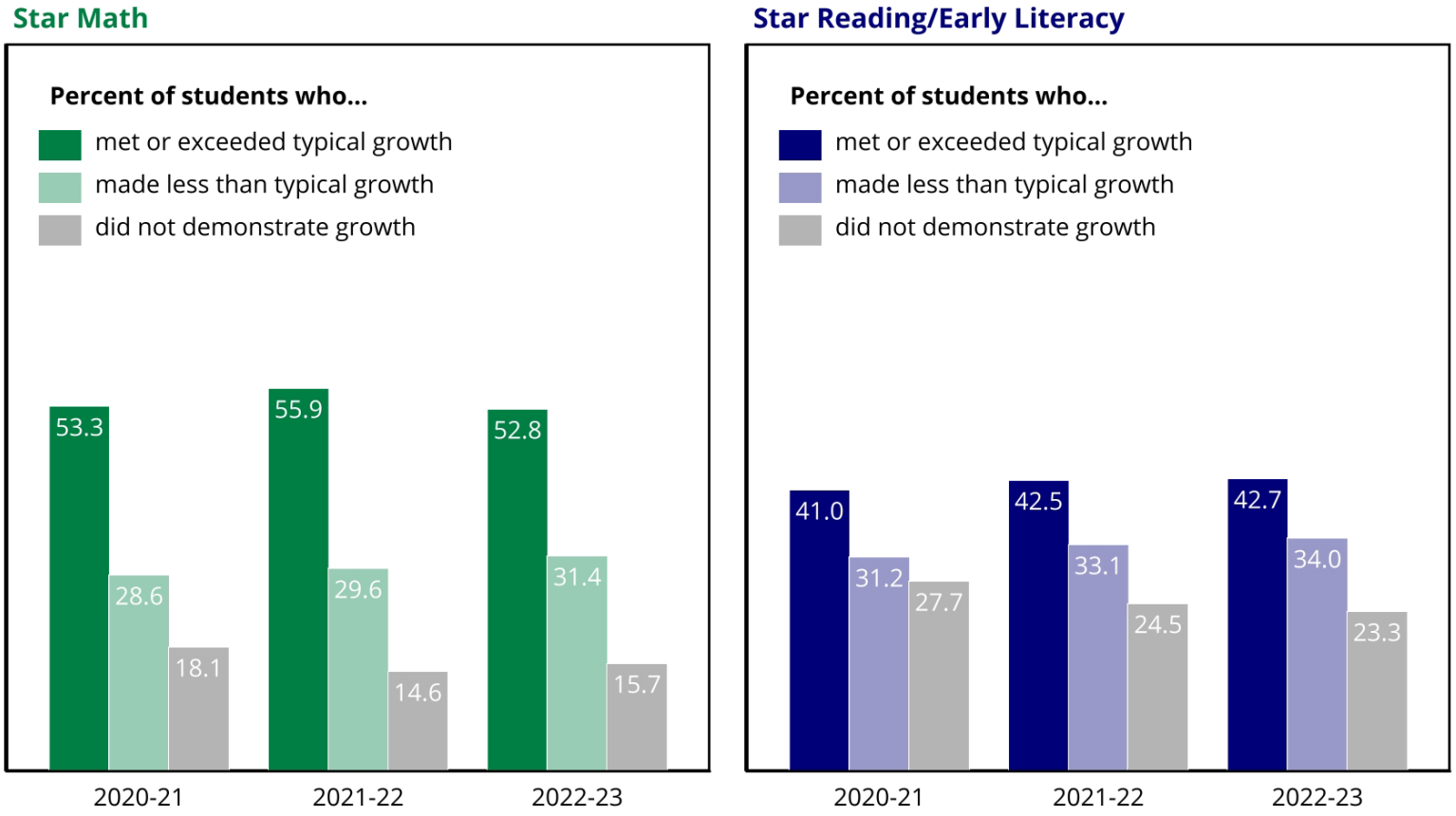
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure A.5.2. Fall-to-Spring Growth Outcomes by School Year (i-Ready)



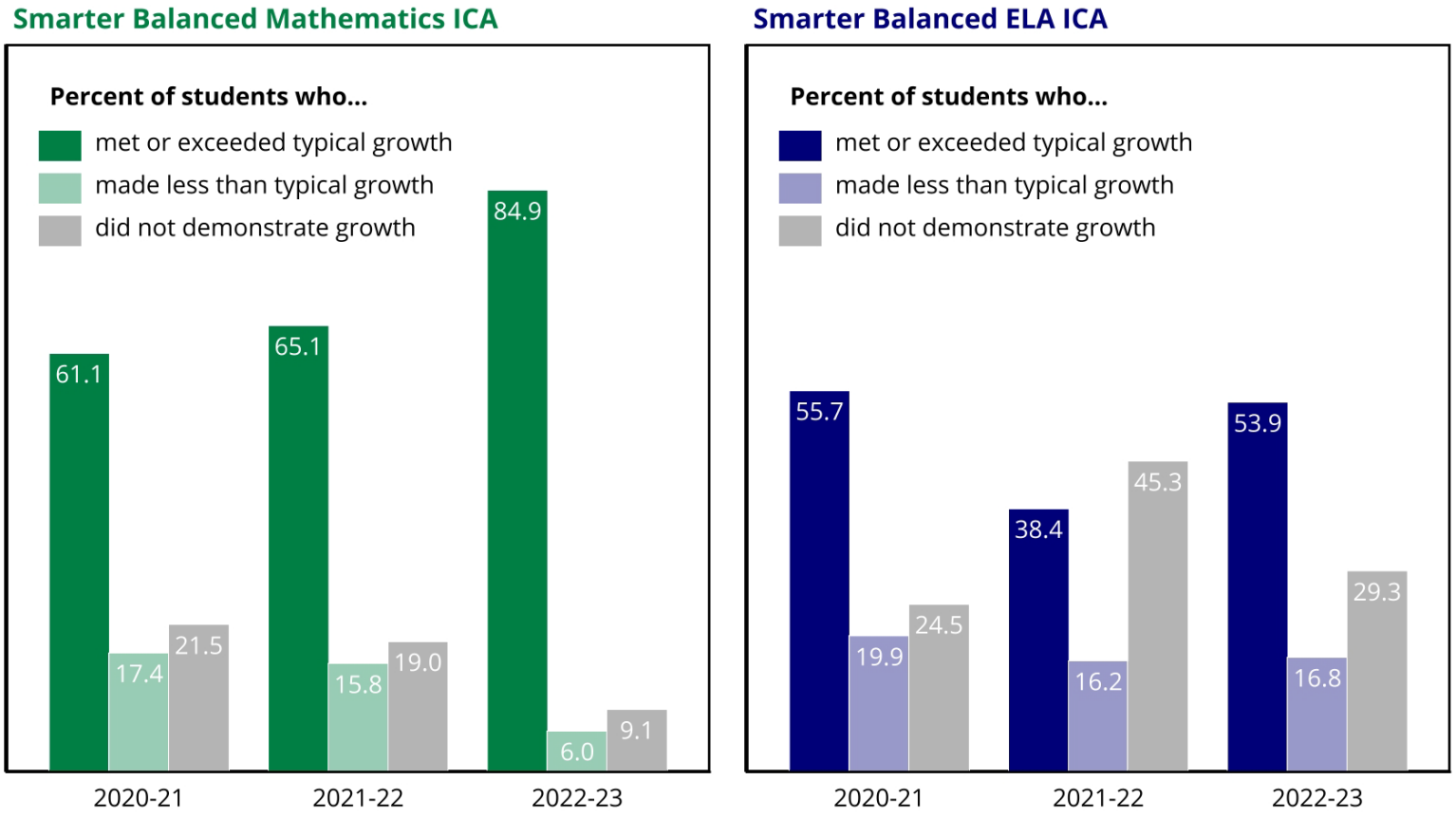
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure A.5.3. Fall-to-Spring Growth Outcomes by School Year (Star 360)



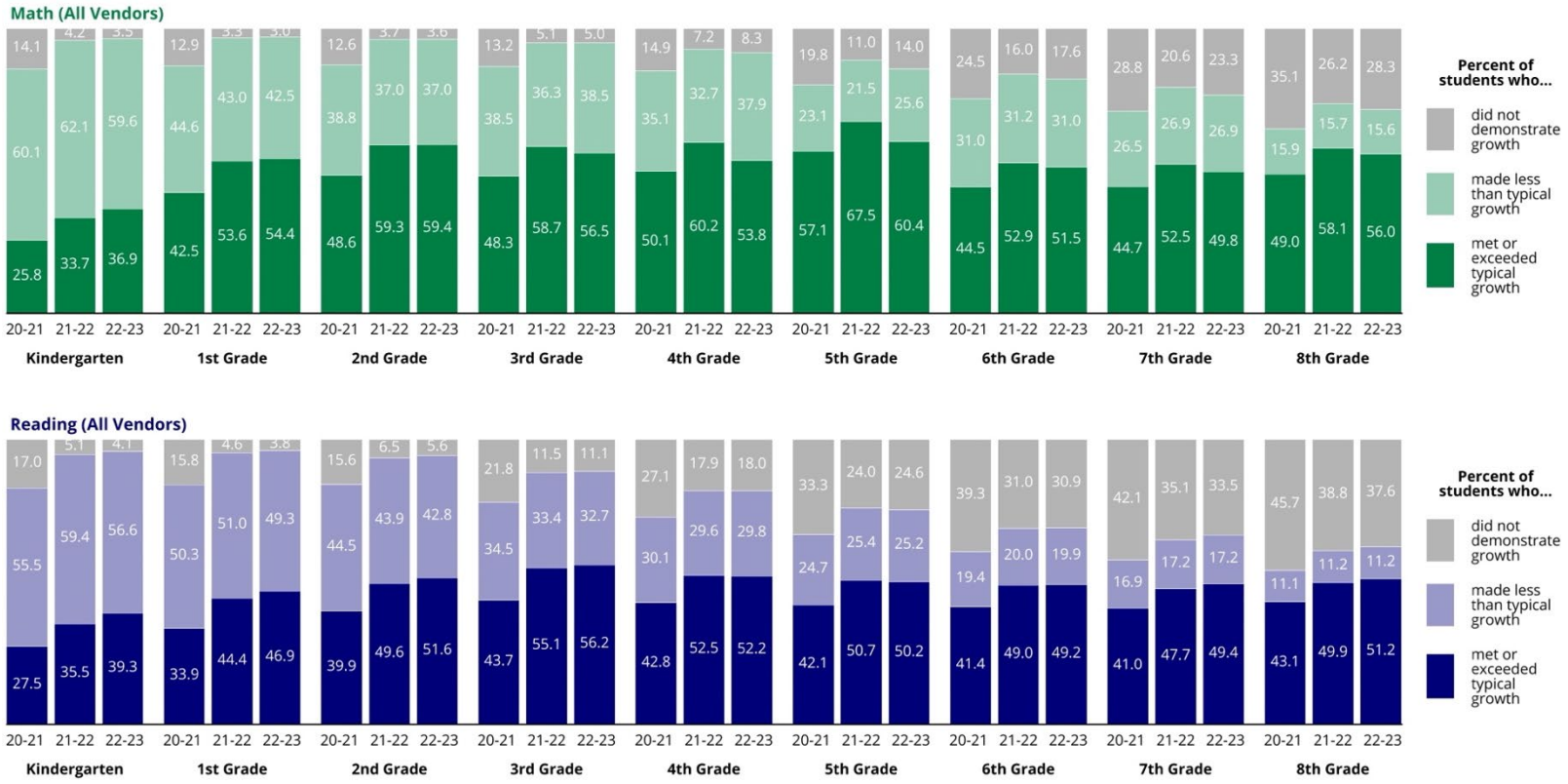
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure A.5.4. Fall-to-Spring Growth Outcomes by School Year (Smarter Balanced ICA)



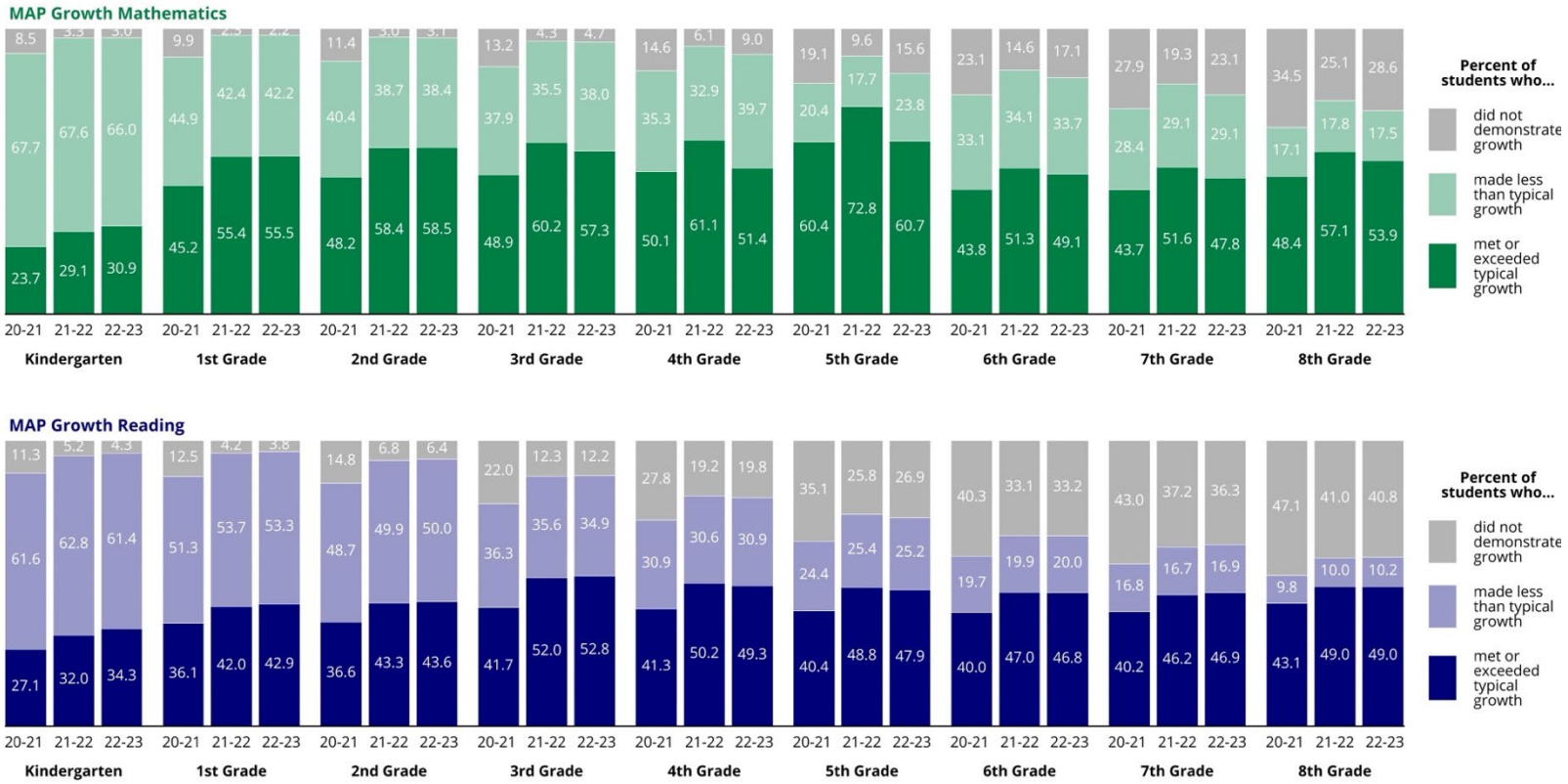
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure A.6.1. Fall-to-Spring Growth Outcomes by School Year and Grade Level (All Vendors)



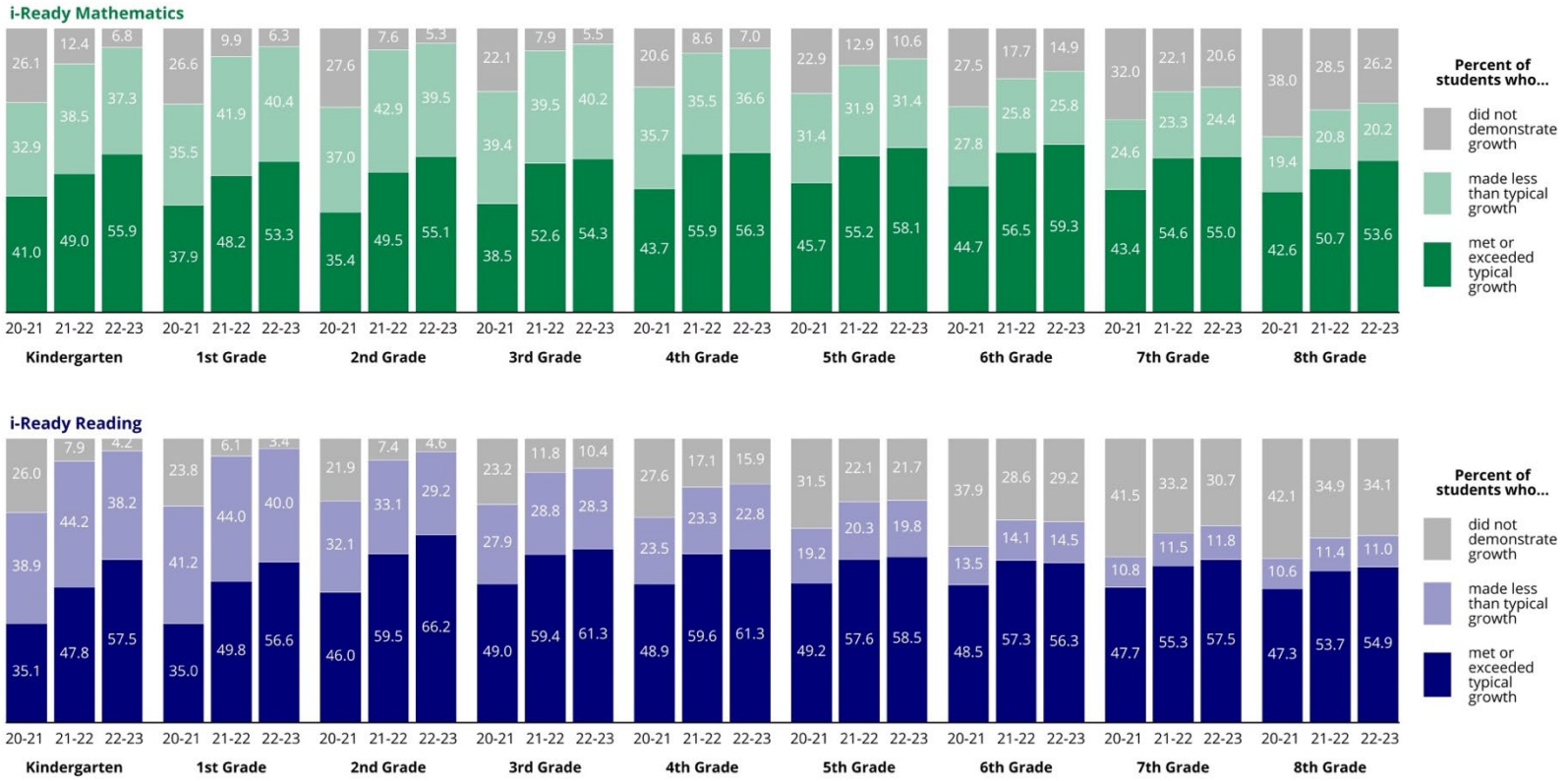
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure A.6.2. Fall-to-Spring Growth Outcomes by School Year and Grade Level (MAP Growth)



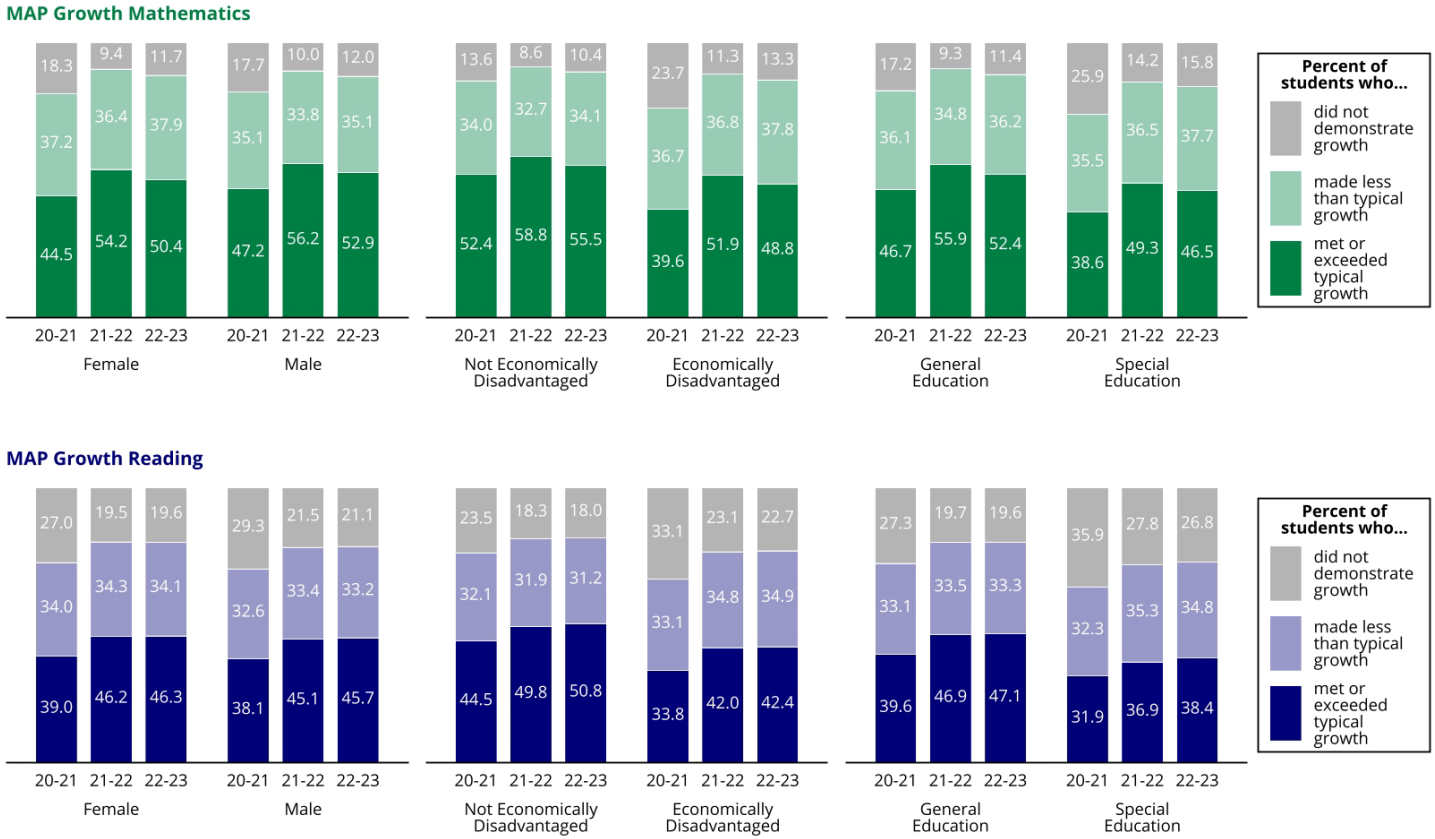
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure A.6.3. Fall-to-Spring Growth Outcomes by School Year and Grade Level (i-Ready)



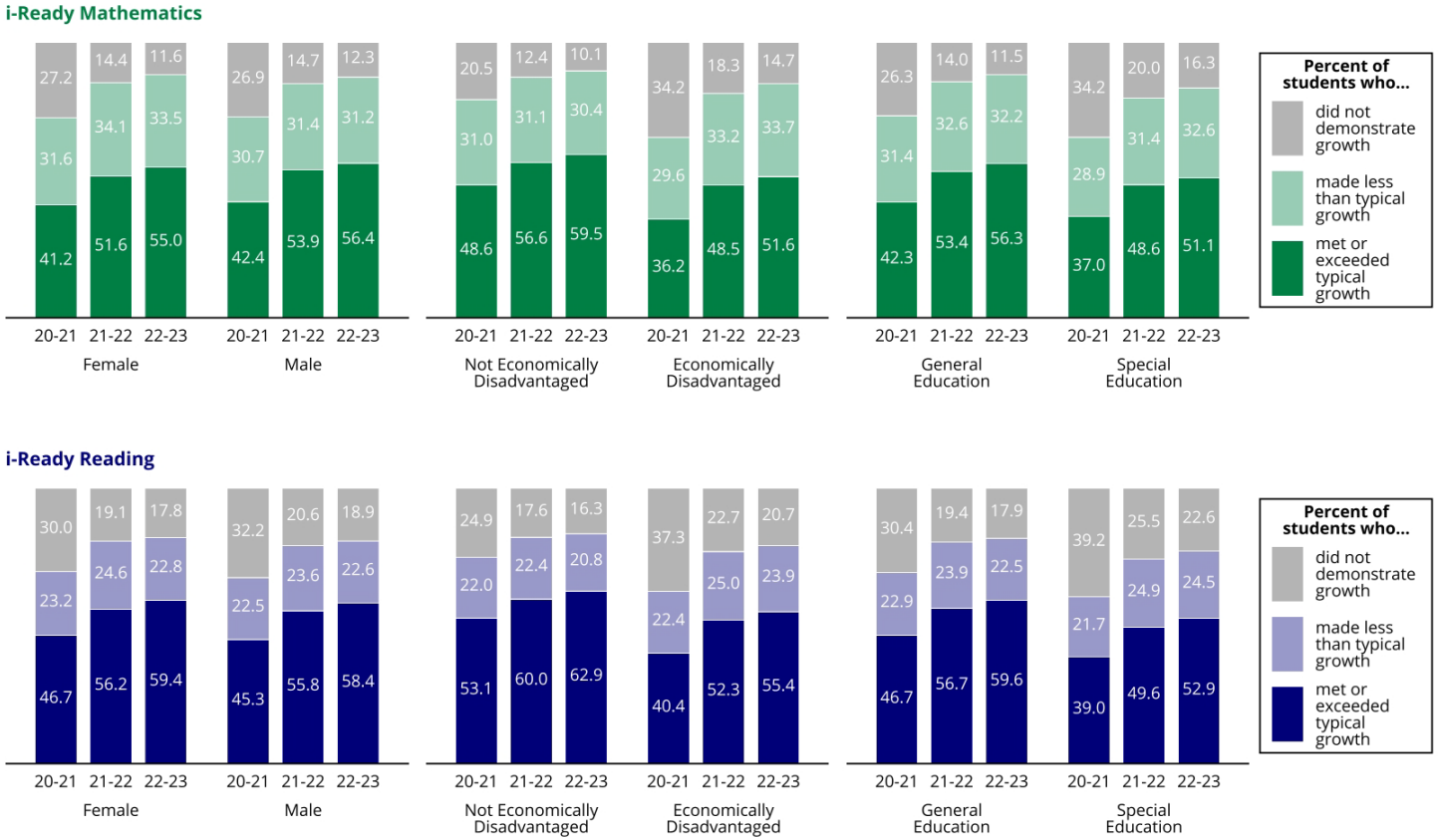
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure A.7.1. Fall-to-Spring Growth Outcomes by School Year and Student Subgroup (MAP Growth)



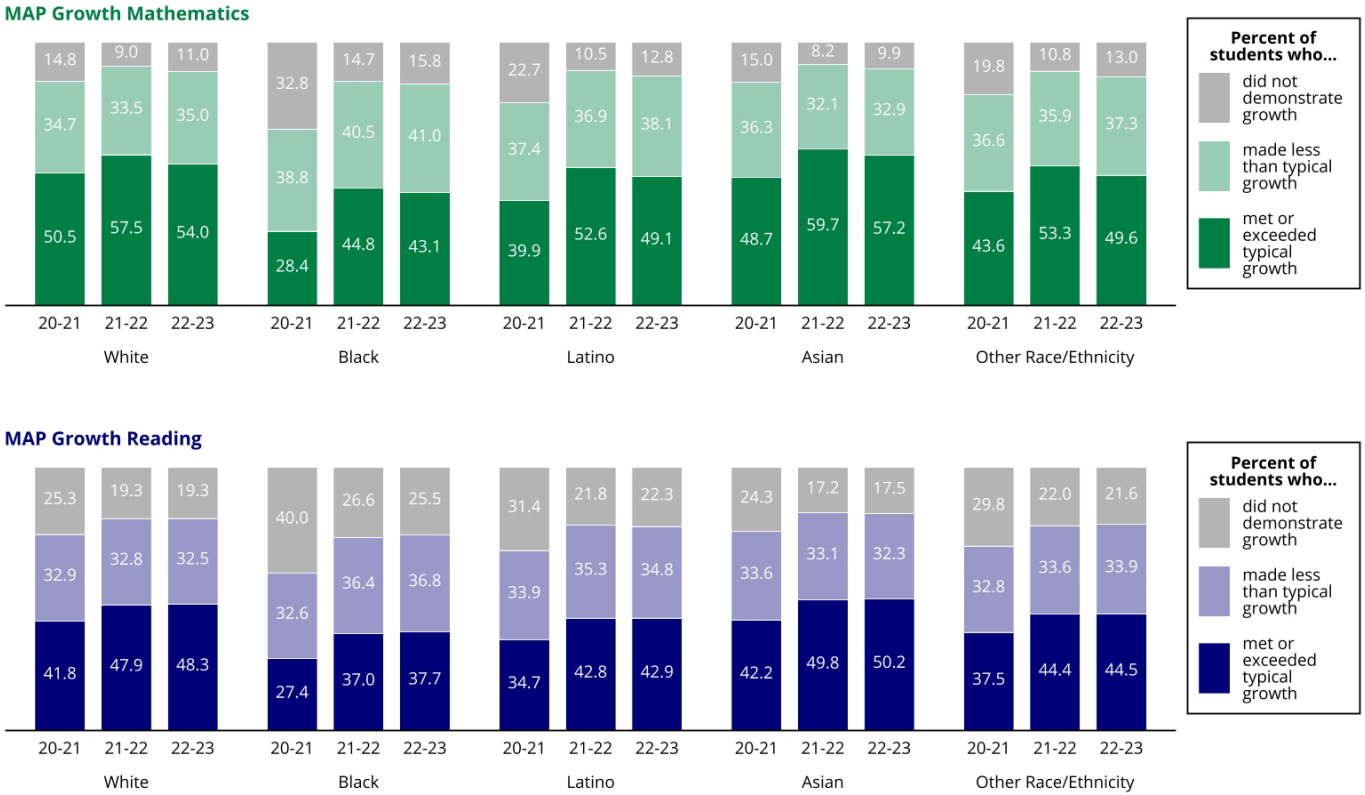
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure A.7.2. Fall-to-Spring Growth Outcomes by School Year and Student Subgroup (i-Ready)



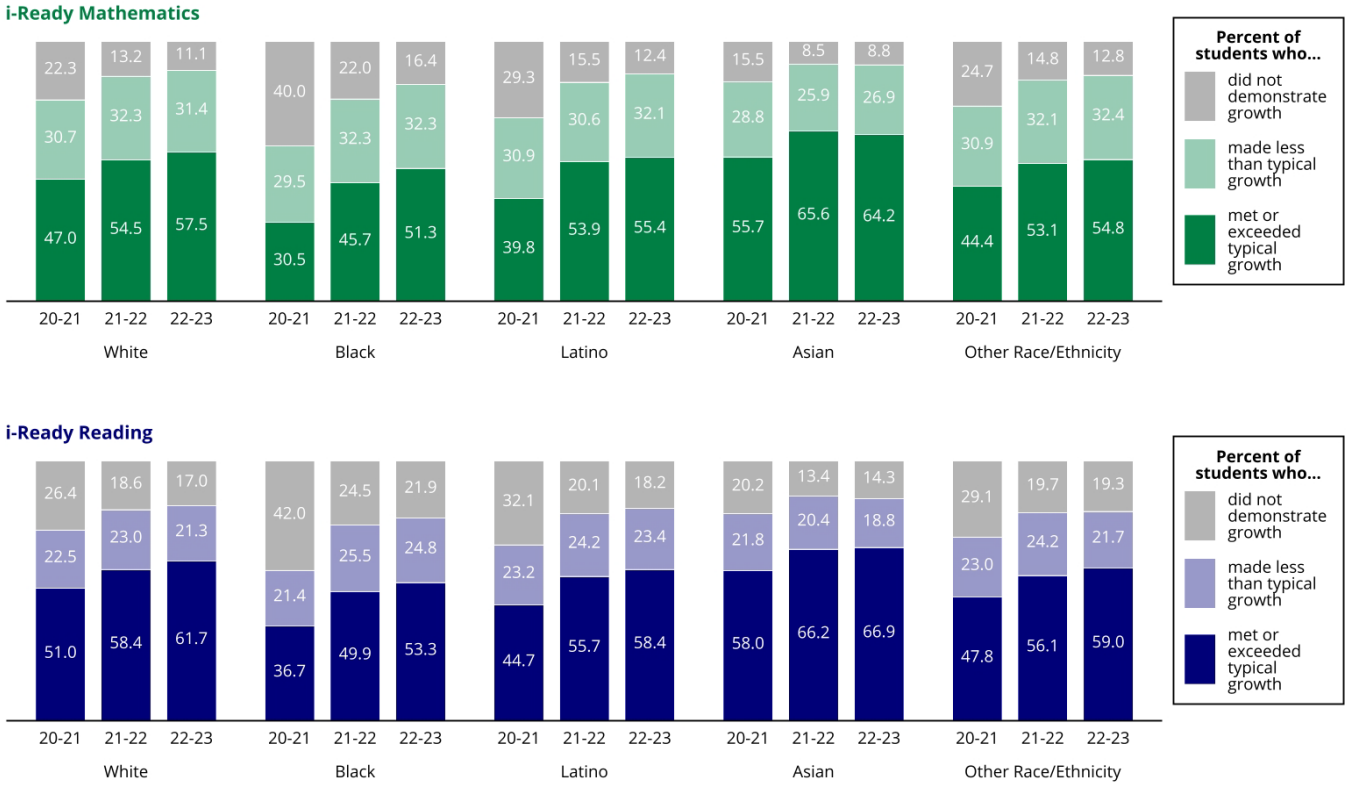
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The percentages may not add up to 100% due to rounding.

Figure A.7.3. Fall-to-Spring Growth Outcomes by School Year and Student Race/Ethnicity (MAP Growth)



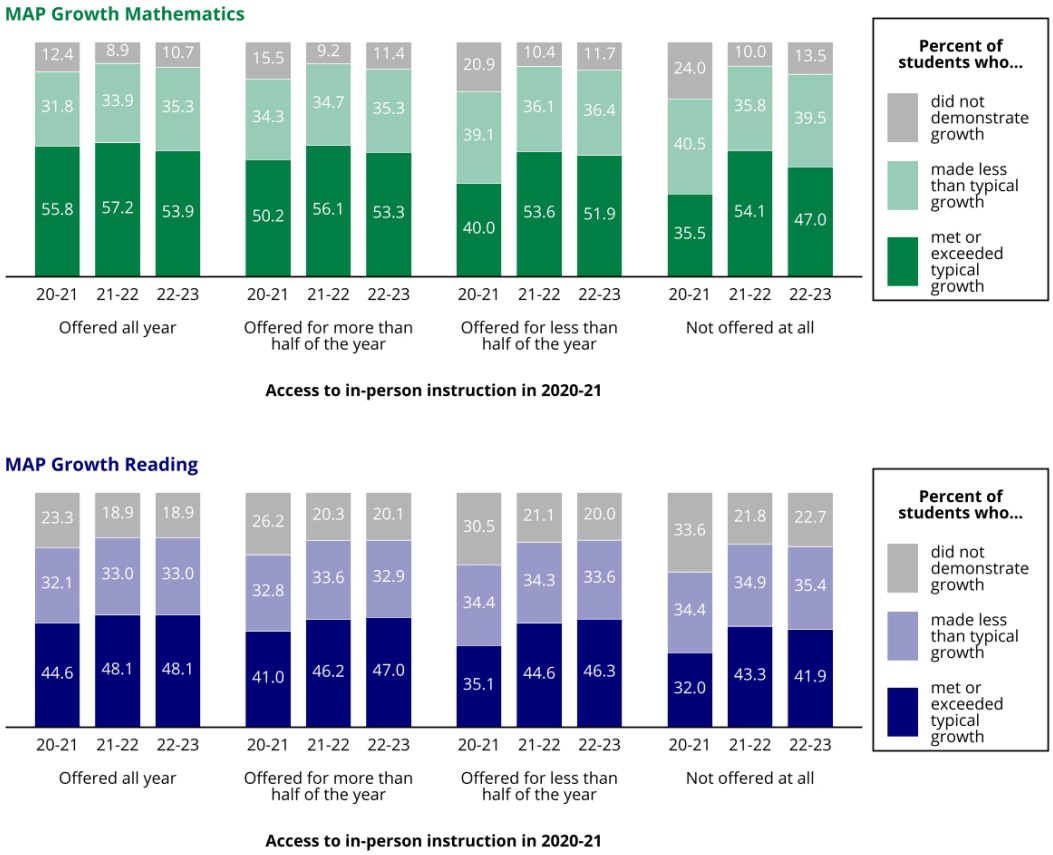
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The "Other Race/Ethnicity" category includes students who are Native American, Hawaiian or Pacific Islander, or two or more races; we cannot show separate bars for each of these groups due to low sample sizes. The percentages may not add up to 100% due to rounding.

Figure A.7.4. Fall-to-Spring Growth Outcomes by School Year and Student Race/Ethnicity (i-Ready)



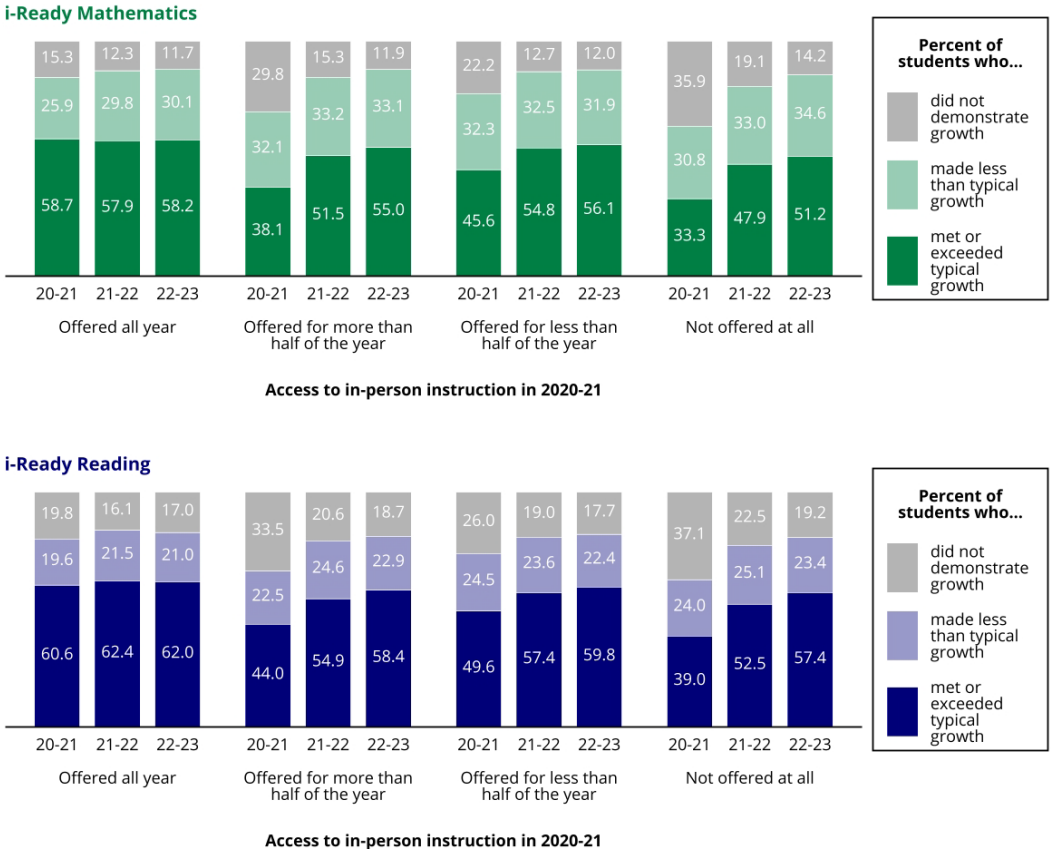
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The "Other Race/Ethnicity" category includes students who are Native American, Hawaiian or Pacific Islander, or two or more races; we cannot show separate bars for each of these groups due to low sample sizes. The percentages may not add up to 100% due to rounding.

Figure A.8.1. Fall-to-Spring Growth Outcomes by School Year and Access to In-Person Instruction in 2020-21 (MAP Growth)



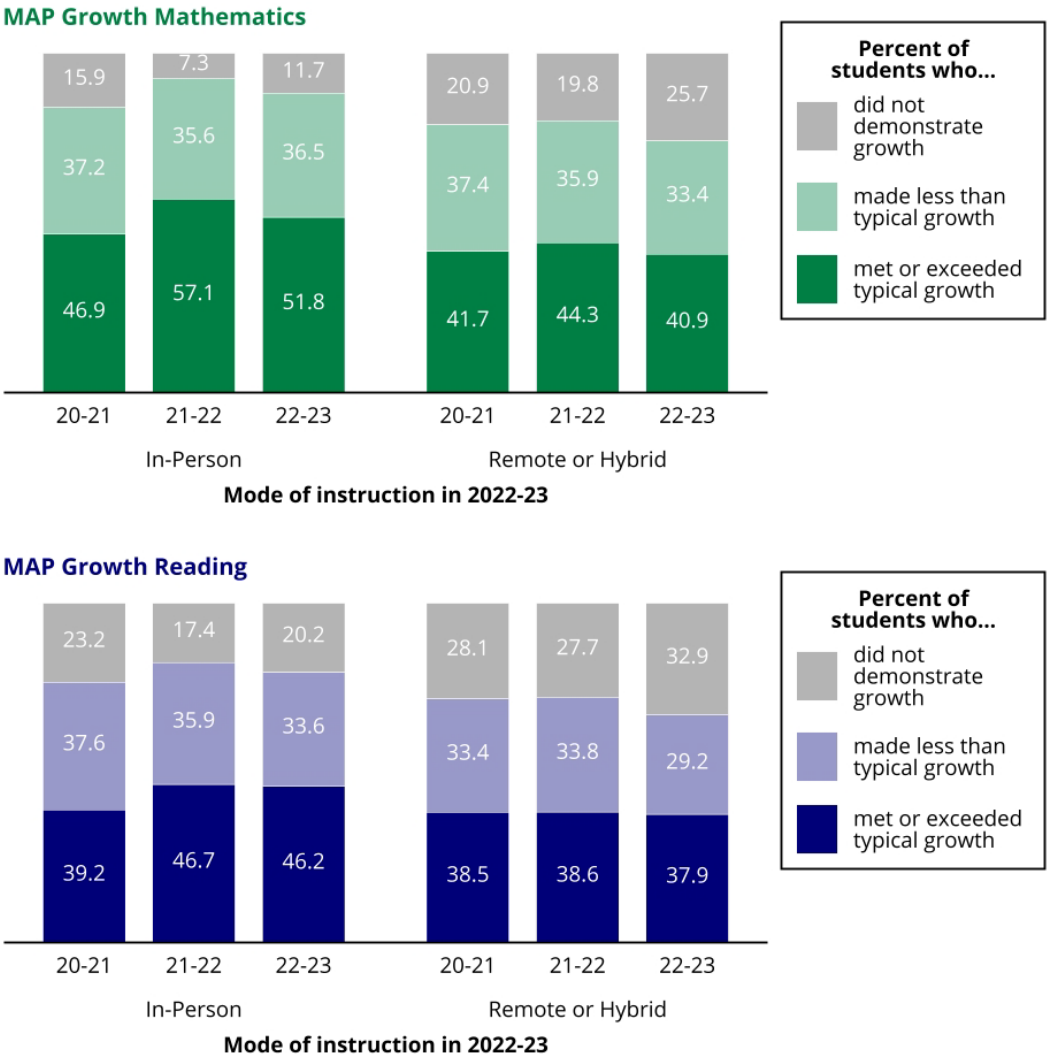
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The modality data from 2020-21 summarize what modes of instruction districts offered during each month of the school year. The percentages may not add up to 100% due to rounding.

Figure A.8.2. Fall-to-Spring Growth Outcomes by School Year and Access to In-Person Instruction in 2020-21 (i-Ready)



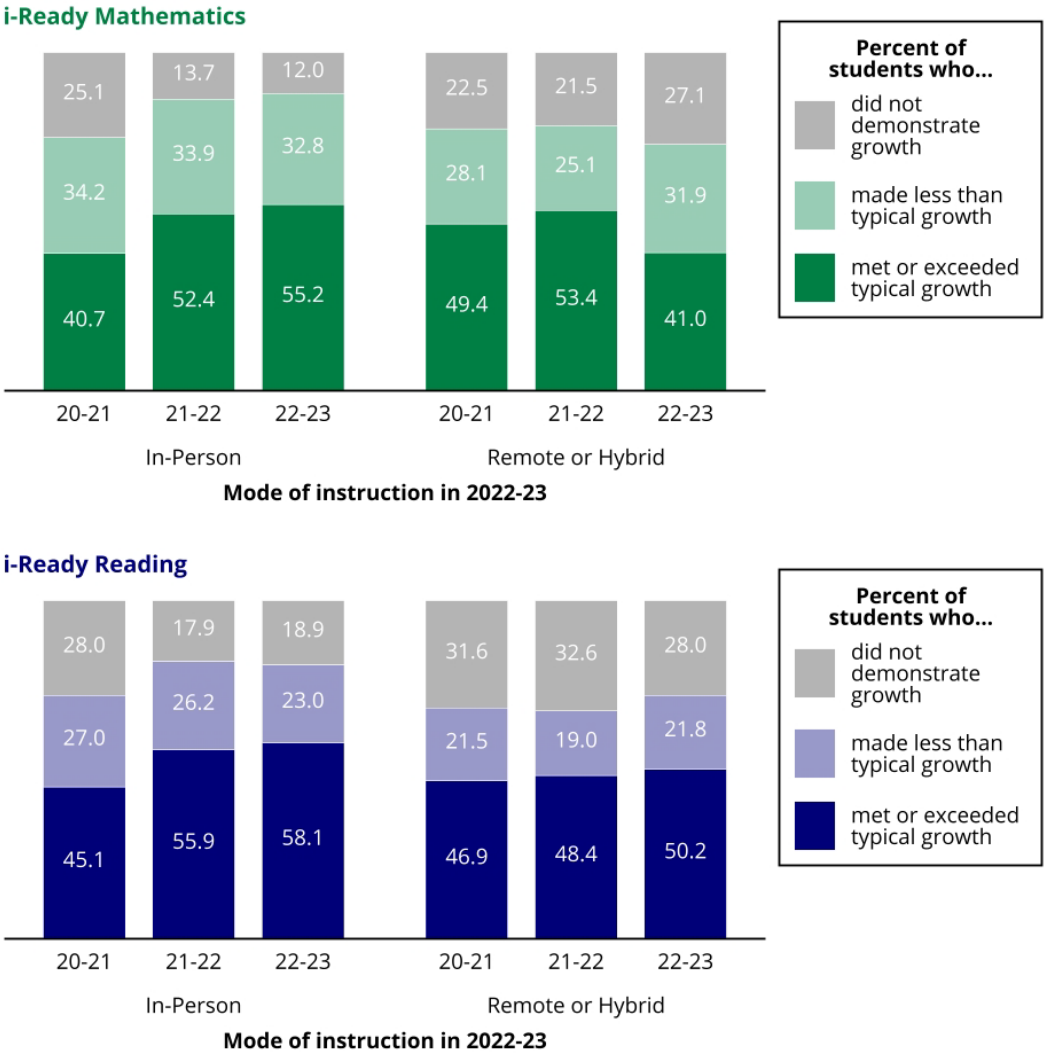
Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The modality data from 2020-21 summarize what modes of instruction districts offered during each month of the school year. The percentages may not add up to 100% due to rounding.

Figure A.8.3. Fall-to-Spring Growth Outcomes by School Year and Mode of Instruction in 2022-23 (MAP Growth)



Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The modality data from 2022-23 summarize what mode of instruction a student received during school year. The percentages may not add up to 100% due to rounding.

Figure A.8.4. Fall-to-Spring Growth Outcomes by School Year and Mode of Instruction in 2022-23 (i-Ready)



Notes: The percentages for each school year only include students with benchmark assessment scores in both the fall and spring. Thresholds for "typical growth" are based on pre-pandemic norms from each assessment provider, which vary depending on the subject area, grade level, and students' initial achievement on their fall benchmark assessment. The modality data from 2022-23 summarize what mode of instruction a student received during school year. The percentages may not add up to 100% due to rounding.

Endnotes

[1] This number is smaller than the 750 districts that provided student-level data for the 2022-23 school year because three of those districts provided data only for grade levels or subjects outside the scope of this report and 13 districts provided data for only fall 2022 but not for spring 2023.
[2] This is largely driven by 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 despite there being 75 districts that used i-Ready.
[3] The results we present in this report are aggregated to the state level. To prevent identifying any individual students from very small subgroups, we do not show results for any cells that represent fewer than ten students.
[4] Some vendors have published new norms for their benchmark assessments since this study began. For comparability across school years and report iterations, we have continued to use the same norms that were in place for each assessment in 2020-21.
[5] Given the differences in the characteristics of students from i-Ready districts relative to the state average (see Section Two of this report), we use the distributions of scale scores from Michigan districts that completed the i-Ready assessments in 2018-19 as comparison points instead of national norms. Michigan students' average i-Ready scores, both before and during pandemic-affected years, are very far below the national norms for these assessments, making it difficult to interpret visual trends when we use these as comparison points. While the Michigan-specific medians from 2018-19 do not represent the same exact groups of students or districts who participated in these assessments in 2020-21 through 2022-23, they represent a more comparable population than the national norming sample.
[6] The Star 360 assessments are scored on two different scales: the Star Enterprise Scale and the Star Unified Scale. The benchmark assessment data available for the reports in this series thus far have included only Enterprise scores for students who take the Star assessments. For this year's report, we used information from the Star 360 technical manuals to convert Enterprise Scale Scores to Unified Scale Scores. This allows us to show trends on both the Star Reading and Star Literacy assessments on the same plot in Figure 3.3.
[7] As we show in Appendix Figure A.2.1, the proficiency rates we estimated from students' benchmark assessment scores and 2022-23 align closely with M-STEP proficiency rates for the same districts in the same years. While we do not know exactly how students would have performed on benchmark assessments in 2018-19, we consider the M-STEP proficiency rates from that year to be a reasonable approximation.
[8] The sample of students who took the Smarter Balanced ICA assessments is the one exception to this pattern. However, it is difficult to interpret trends over time for this

sample because of changes in the number and characteristics of students who participated in these assessments each year. While some districts used the Smarter Balanced assessments across the three years, sometimes districts administered the Smarter Balanced IABs, which cover only a small number of subtopics, to students in some grade levels and administered the Smarter Balanced ICAs, which cover a broader range of math or ELA content, to students in other grades. We can include only the ICAs in our analysis, as the sub-area assessments are not comparable to the broader ICAs or the other MDE-approved benchmark assessments. For this reason, we omit Smarter Balanced ICA results from some of our analyses.

[9] We note that this definition of "typical growth" or "growth targets" differs in meaning from the way practitioners use similar terms, such as "growth goals," in the classroom. In classroom contexts, teachers likely set student growth goals or targets that represent what the teacher expects an individual student to achieve in a given period. This is different from the targets for "typical growth" that we use in this report, which indicate *the median growth that students with similar prior scores achieved before the pandemic*.

[10] Although we include data from all four benchmark assessment vendors in our overall growth analyses, we provide appendix figures with vendor-specific subgroup breakdowns only for MAP Growth and i-Ready. This is because there are too few students in the Star 360 and Smarter Balanced ICA samples to make meaningful comparisons across subgroups.