TECHNICAL BRIEF

Technical appendix for:
Learning during COVID-19: Initial findings on students' reading and math achievement and growth

November 2020

Megan Kuhfeld, Erik Ruzek, Angela Johnson, Beth Tarasawa, and Karyn Lewis

Suggested citation: Kuhfeld, M., Ruzek, E., Johnson, A., Tarasawa, B., & Lewis, K. (2020). Technical appendix for: Learning during COVID-19: Initial findings on students' reading and math achievement and growth. NWEA.

Table of Contents

List of Tables

List of Figures

1. Introduction

The SARS-CoV-2 (COVID-19) pandemic has disrupted traditional forms of education and continues to create challenges for K-12 school systems in the United States and the students they serve. These unique educational challenges have occurred at the same time as a major economic downturn, job losses, widespread protests over racial injustice, a presidential election, and the continuing tangible health threat posed by COVID-19. The public health, education, and economic damages inflicted by COVID-19 are likely to exacerbate long-standing inequities disproportionately affecting Black, Latinx, Native American students, English Learners, and students with disabilities. Beyond educating our children, schools also play important roles in providing essential services across many communities, such as after-school programs, pre-K offerings, counseling, and meal programs. Districts attempted to transition academic and non-academic activities to remote settings in spring of 2020, and many continue to do so through the fall of 2020, with varying degrees of success.

Policy makers and educational leaders have the unenviable responsibility of making difficult decisions well into the 2020-21 school year and beyond. Now, more than ever, we need data to inform evidence-based policies to support our students, teachers, and families on the path to recovery. To help educators and policymakers better prepare for the potential magnitude of the impact of COVID-19 on learning, NWEA previously released a set of projections[i] of the potential academic impact of COVID-19 disruptions modeled on well-documented summer learning loss estimates. NWEA researchers and collaborators subsequently published a more sophisticated set of projections[ii] that were informed by research not only on summer learning loss but also by research on student absenteeism. The projections considered multiple scenarios: (a) typical learning (where students would be expected to be in a normal school year), (b) partial absenteeism (assuming students received half of their normal instruction in the spring), and (c) COVID-19 Slide (where the spring school closures were assumed to operate like a typical summer break from school). Other organizations released projections, including CREDO's state-level projections[iii] and McKinsey's projections[iv] that attempted to break down the loss by different student demographic groups and under different assumptions about when students would return to in-class instruction. In general, all available studies projected that there would be detrimental impacts of COVID-19 school closures on student learning.

As students have returned to school in the fall, schools are working hard to understand just how much the school closures and disruptions have set students back and how they can best support students who were particularly hard-hit by the shutdown. With actual fall data now in hand, we can move beyond forecasting what might happen to begin to describe what did happen. We recently released a research brief[v] summarizing our findings of how school shutdowns impacted student achievement at the start of the 2020-21 school year. This research was conducted with fall 2020 data on the NWEA MAP Growth test taken by millions of students in grades 3-8 in over 8,000 schools across the U.S. The scope of these data is unparalleled, providing a national perspective on student progress since COVID-19 school closures in March 2020.

The purpose of this technical brief is to share more detailed results and to describe more fully the sample and methods used to address three primary research questions in that study:

1. How did students perform in fall 2020 relative to a typical school year (specifically, fall 2019)?
2. How has student growth changed since schools physically closed in March 2020?
3. How did observed fall 2020 achievement compare to NWEA's projected scenarios?[ii]

2. Data

Sample

The data for the study are from the NWEA anonymized longitudinal student achievement database. School districts use NWEA MAP Growth assessments to monitor elementary and secondary students' reading and math growth throughout the school year, with assessments typically administered in the fall, winter, and spring. The NWEA data also include demographic information, including student race/ethnicity, gender, and age at assessment, though an indicator of student-level socioeconomic status is not available. A set of school-level characteristics, including school-level free or reduced priced lunch (FRPL) eligibility and school location (urban, rural, town, or suburb), was obtained from a data file produced by the Stanford Education Data Archive (SEDA) version 3.0.[vi]

In total, slightly under 7 million students in 20,000 public and private schools took MAP Growth reading and math assessments in fall 2020. From that initial sample of test takers, we further limited to public schools that had previously tested in the 2019-20 school year, which resulted in a sample of 4.4 million students. To best address our research questions, we defined different analytic samples that sought reduce the degree to which changes in the schools and students in the sample tested drove changes in the observed results. A cross-sectional sample was created for the first research question (which compared grade 3-8 students in fall 2020 with their same-grade peers in fall 2019) and then a longitudinal sample was created for the remaining two research questions (which followed students between the 2019-20 and 2020-21 school years). The first sample consisted of approximately 9,000 schools that tested students in both fall 2019 and fall 2020 and was used for the cross-sectional comparison of students in grades 3-8 in fall 2020 with their same-grade peers in fall 2019.[1] The second sample further restricted to students who consistently tested in fall 2019 (August to November 2019), winter 2020 (January through early March 2020), and fall 2020 (August to October 2020). More details on each are provided below.

Analytic Sample 1. To answer the first research question around within-grade comparisons between fall 2019 and fall 2020, we limited our sample of schools to a consistent set of U.S. public schools that tested at least ten students in a given grade in both fall 2019 and fall 2020. This sample restriction reduces the degree to which changes in the NWEA partner base drive the observed results (see our attrition analysis brief[vii] for an examination of attrition patterns within this sample of schools). This sample restriction guards against the competing explanation that any differences we observe in achievement over time are potentially driven by systematic differences between schools that did and did not test students in fall 2020. In total, the first analytic sample contained 3,267,867 unique students in 8,961 schools in reading and 3,249,883 unique students in 8,874 schools in math. Table 1 provides a comparison of the students in the

[1] Results from our comparability analysis[viii] of remote and in-person testing suggest that the remote testing experience is consistent with in-person testing for students in grades 3-8, but may qualitatively differ for the youngest students. Please see the full comparability analysis for more information.

fall 2019 and fall 2020 analytic sample by grade in reading, and Table 2 provides the same information for math. Overall, the samples of students that tested in 2019 and of same-grade students that tested in fall 2020 were very similar in terms of gender and race/ethnicity,[2] though the number of students tested in each grade was consistently larger in fall 2019.

Analytic Sample 2. For our second and third research questions, our sample included a longitudinal cohort of students that were followed across fall 2019, winter 2020, and fall 2020. Specifically, we followed a cohort of students from third grade in 2019-20 to fourth grade in fall 2020, a separate cohort from fourth grade in 2019-20 to fifth grade in fall 2020, and so on. To be included in this sample, students were required to have an observed test score in each of the three time points mentioned above (approximately 30 percent of the students with 2019-20 data). Table 3 presents the demographic characteristics for students in the second analytic sample. In total, 1,700,798 students in 8,067 schools were included in our longitudinal analytic sample in reading, and 1,854,525 students in 9,492 schools in math. As a reference distribution for estimating "typical" growth during a pre-COVID period, we relied on the prior cohorts of students who tested in winter 2019 and fall 2019 within each grade pair (e.g., grades 3-4, grades 4-5, etc.).

The schools in our samples differ in some important ways from the overall set of U.S. public schools serving grades 3-8 (see Table 4). The schools in our sample comprise approximately one in ten U.S. public schools, and the distribution of schools across various locales (urban, suburban, rural, and town) closely matches the population of schools in the U.S. However, our sample reflects a higher-than-average percentage of White students and slightly lower-than-average percentage of students eligible for free or reduced-price lunch (FRPL).

Measure of achievement

Student test scores from NWEA MAP Growth reading and math assessments were used in this study. MAP Growth is a computer adaptive test that precisely measures achievement even for students above or below grade level and is vertically scaled to allow for the estimation of gains across time. The MAP Growth assessments are typically administered three times a year (fall, winter, and spring) and are aligned to state content standards. Test scores are reported on the RIT (Rasch unIT) scale, which is a linear transformation of the logit scale units from the Rasch item response theory model.

In this study, we used both students' RIT scores and their achievement percentile scores. Achievement percentile scores were calculated using the NWEA 2020 MAP Growth norms.[ix] Since MAP Growth can be taken at any point during the school year, the MAP Growth achievement norms condition on each student's grade, subject, and instructional week of testing (i.e., the week in the school calendar in which a student tested). Instructional weeks were calculated for each student based on their school start date and the individual student's testing dates (for more details on the calculation of instructional weeks, see the norms study[ix]). Within each grade and subject, let Y_{it} be a student *i*'s RIT score at instructional week *t*. The predicted mean $(\hat{Y}_t)$ and standard deviation $\left(SD(Y_t)\right)$ for a given grade/subject/instructional week

[2] Analytic Sample 1, in drawing only on schools that tested in both fall 2019 and fall 2020 is similar to the sample for the second research question in the attrition brief (Table 2 in our attrition analysis[vii]); however it differed in that the attrition brief included children who were in grades K-7 in fall 2019 whereas here and in the main brief we included students who were in grades 2-7 in fall 2019.

combination were pre-calculated based on the NWEA norms model (see Chapter 4 of the norms report). Based on these values, we calculated a standardized estimate of the student's RIT score:

$$z(\mathrm{Y_{it}}) = \frac{(\mathrm{Y_{it}} - \widehat{\mathrm{Y}}_{\mathrm{t}})}{\mathrm{SD}(\mathrm{Y_t})}.$$

From the standardized score, we calculated the score percentile (e.g., the proportion of the distribution that the student scored as well as, or better than):

$$ps(Y_{it}) = Pr(Y_{it} \leq y_t) = \int_{-\infty}^{y_t} \phi(z)dz,$$

where $\phi(z)$ represents the probability density function. The student normative percentile used in this study was scaled to range from 1 to 99:

$$\mathrm{Perc} = \ 100 \times p_s(Y_{it}).$$

3. Methods

RQ1: How are students doing in fall 2020 relative to a typical school year?

To address the first research question, we calculated the median student achievement percentile in fall 2019 and fall 2020 for each grade level and subject. In addition, results were broken out further by student race/ethnicity, school poverty level, and urbanicity. School poverty level was calculated based on the reported percentage of students eligible for FRPL. Due to recent changes in reporting practices on FRPL, we] chose to use the school-level percentage FRPL-eligibility variable from SEDA, which is weighted average of a school's percentage FRPL-eligibility from 2009 to 2016 (see the Stanford Education Data Archive[vi] for more details). We classified schools into three poverty levels: (a) less than 25% FRPL eligibility, (b) 25-75% FRPL eligibility, and (c) greater than 75% FRPL eligibility.

RQ2: Have students made academic gains since schools shut down in March 2020?

We measured growth in this study in two ways. First, we calculated a difference score by subtracting each student's winter 2020 score from their fall 2020 score. We similarly did this for the cohort of students from the prior year (winter 2019 to fall 2019). The two difference score distributions were then compared. Difference scores helps us understand and describe students' raw growth. However, students grow at different rates as they age and only examining raw growth may mask important shifts for older students. Thus, we also looked at changes in normative achievement status by binning students' achievement into quintiles (e.g., 1-20th percentile, 21-40th percentile, 41-60th percentile, 61-80th percentile, and 81-99th percentile) using their calculated percentile ranks separately within the winter and fall terms. We then calculated the percentage of students who stayed in the same quintile of the distribution in winter 2020 and fall 2020 ("Maintainers"), the percentage that moved to a higher quintile in fall 2020 relative to their winter 2020 quintile ("Gainers"), and the percentage that moved down a quintile in fall 2020 ("Sliders"). An example of the assignment of students to different groups is given in Table 5 based on the math test scores for students moving from grade 3 to 4 between winter 2020 and

fall 2020. We used the same procedure for the previous longitudinal cohort of students who tested in winter 2019 and fall 2019.

RQ3: How do student test scores in fall 2020 compare to the NWEA COVID Slide projections?

NWEA released a set of projections[i] in spring 2020 for how much learning loss students might experience because of COVID-19 spring 2020 school closures and, subsequently, with collaborators, refined these projections.[ii] In this study, we focused on three of the projected scenarios: (a) typical learning (where students were projected to learn at a rate consistent with a normal school year— a best case scenario in which there was no effect from COVID-19), (b) partial absenteeism (which assumed students received half of their normal instruction in spring 2020), and (c) COVID-19 Slide (where the spring 2020 school closures were assumed to operate like a typical summer break from school, applying typical summer learning loss rates across the extended spring closure).To compare the observed data to our projections, we calculated the average test score for students who tested within specific timeframes in fall 2019, winter 2020, and fall 2020 and overlaid those means on our projected academic growth trajectories. Given the wide variability in testing windows across districts (for more detail, see the 2020 NWEA norms report[ix] Table 2.4), we decided to only plot test scores from time points with high data coverage. Specifically, we pooled scores from students who tested between weeks 3 and 5 of the 2019-20 school year to calculate the fall 2019 average test score; students testing in weeks 19 to 21 for the winter 2020 average; and students testing in weeks 3 to 5 of the 2020-21 school year to represent fall 2020.

4. Results

RQ1: How are students doing in fall 2020 relative to a typical school year?

Compared to same-grade students in fall 2019, students in grades 3-8 in fall 2020 generally had similar percentile scores in reading (top panel of Figure 1). However, percentiles for math were considerably lower, on average, for students this year as compared to similar-grade students last year (bottom panel of Figure 1).[3] Figure 2 further shows the percentile differences across ethnic/racial groups. While most groups showed stable performance (or even slight improvements over fall 2019), small drops were observed in reading for Black and Hispanic students in upper elementary grades. In math, fall 2020 students in all ethnic-racial groups had lower percentile scores compared to same-grade students tested in fall 2019. These latter results should be considered in the context of the nontrivial numbers of predominantly ethnic-racial minority students that were present in the fall 2019 data but not in the fall 2020 data from schools that tested in both time periods (see the accompanying attrition brief[vii]). Because more students in these demographic groups did not test in fall 2020, the differences reported here are likely to be underestimated.

Figure 3 breaks down the percentile differences in fall 2019 and fall 2020 by school poverty level. In reading, students in high-poverty schools showed a lower median percentile rank in grades 3-6 in fall 2020 compared to fall 2019, whereas students in low-poverty schools performed similarly across the two terms. In math, sizable drops were observed across all

[3] The fact that the students in this sample were consistently above the 50th percentile in fall 2019 indicates that they are higher achieving than the national population on which the NWEA norms are based.

school poverty levels in grades 3-6. In grades 7 and 8, percentile drops in math were smaller in high-poverty schools relative to low- and mid-poverty schools. Figure 4 presents the percentile differences in fall 2019 and fall 2020 by school location (urban, town, rural, or suburban). Results did not notably vary by school urbanicity.

Sensitivity analysis – Providing informed bounds to account for students not testing in fall 2020
A quarter of the students who tested in fall 2019 did not test in fall 2020 and thus were not included in our analyses. To test the sensitivity of RQ1 findings to the inclusion of these missing students, we approximated their fall 2020 test scores using four potential scenarios. The four achievement percentiles assigned to students not testing in fall 2020 included: (1) the missing students' fall 2019 percentiles; (2) their fall 2019 percentiles adjusted downwards by subtracting the median percentile rank drop in the observed sample of students (1 point in reading, 5 points in math); (3) the median fall 2020 percentile of their same grade-gender-race peers; (4) the median fall 2020 percentile of their same grade-gender-race peers adjusted downwards. For each of the four potential scenarios, we filled in the approximated percentile rank for the missing students to the fall 2020 data and reran the analyses for RQ1 to get four sets of results to compare to our analytic sample without the missing students. That is to say, we estimated the change in median percentile rank between fall 2019 and fall 2020 under each of the four scenarios and compared results to the estimates based on only observed students. The lowest of the five percentile scores forms the lower bound, or a worst-case scenario for achievement in fall 2020; the highest of the five forms the upper bound, or a best-case scenario. Table 6 shows the original RQ1 estimates of the differences between fall 2019 and fall 2020 median percentile scores of same-grade students ("observed difference") and their corresponding lower and upper bounds. For math, even our best-case scenario estimates suggested that fall 2020 achievement scores were between 4 and 9 percentile points lower than same-grade peers in fall 2019. For reading, our worst-case scenario estimates suggested a small drop (1-3 percentile points) in percentile rank, and our best-case scenario estimates suggested a small increase for most grades.

RQ2: Have students made academic gains since schools shut down in March 2020?

Figure 5 shows the distribution of within-student growth from winter 2020 to fall 2020, compared to students making parallel grade transitions between winter 2019 and fall 2019. The black dashed line represents zero growth (winter and fall test scores were equivalent), the red line shows median growth from winter 2019 to fall 2019, and the blue line shows median growth from winter 2020 to fall 2020. We observed that over 50 percent of students showed at least some growth in math and reading between winter 2020 and fall 2020, with the lone exception of students moving from fifth to sixth grade in math. The 2019 and 2020 reading distributions largely overlapped, suggesting the distribution of growth since winter 2020 (when the COVID-19 pandemic began) was parallel to a "typical" year. In contrast, a smaller proportion of students showed positive math growth in the 2020 period relative to 2019 for all grade levels. It is also worth noting that while technically a small majority of students did make some gains in math, losses were evident for almost half of the students in many grades. That was not the case for 2019, in which less than 40% of students lost ground in most grades.

Figure 6 provides a different way of examining whether students maintained or lost ground since the COVID-19 pandemic started. Instead of comparing RIT score gains/losses between winter and fall, Figure 6 illustrates changes in students' relative rank since the beginning of the pandemic compared to the prior year. Specifically, this figure presents the percentage of

students by grade level that stayed within the same achievement quintile (Maintainers), shifted upwards (Gainers), or downwards (Sliders) from winter to fall. In most grades in math, nearly twice as many students moved down at least one quintile in 2020 than in 2019 (see lower panel of Figure 6). For example, while 16.6% of students in grade 4 in 2019 were Sliders in math, that percentage more than doubled (37.7%) in 2020. In reading, in contrast, there was no dramatic shift: the percentage of Gainers, Maintainers, and Sliders was similar in 2019 and 2020 (see upper panel of Figure 6).

RQ3: How do student test scores in fall 2020 compare to the NWEA COVID Slide projections?

Figure 7 overlays labeled circles representing the average observed scores of students in our sample during the 2019-20 and 2020-2021 school years onto a graph of our prior projections[ii] of the COVID slide. Though represented as a single circle for each grade for each test window (fall 2019, winter 2019, and fall 2020), the means are from all students in this study's sample who tested at all time points within that term's testing window. It is worth noting is that students in this sample were higher achieving relative to the broader sample taking the MAP Growth test prior to COVID disruptions in fall 2019 (so, the circles representing fall 2019 achievement in each grade in both math and reading are above the typical learning projection). The solid and dashed lines shown in the 2019-20 school year represent the our previously published projections.[ii] Because these scenarios only projected to the start of the 2020-2021 school year (e.g., September 1, 2020), we extrapolated academic growth assuming typical grade-level growth (based on NWEA MAP Growth norms) to extend the projections through fall 2020 for all three scenarios during the 2020-21 school year. As a result, the three lines displayed are parallel within each grade across fall 2020. In reading, average fall 2020 reading test scores were consistent with the typical learning scenario (see upper panel of Figure 7). In fact, the sixth- and eighth-grade averages at all three time points are at or above the typical reading growth trajectory, reflecting the fact that the analytic sample is performing above the 50th percentile prior to the start of the pandemic and were aligned with the typical-learning scenario through fall 2020.

In math, in contrast, (see lower panel of Figure 7), students tested in fall 2020 had, on average, scores that ranged from on the typical growth trajectory (e.g., eighth-graders) to below the typical growth line and into the partial absenteeism growth scenario (e.g., fourth- and sixth-graders). In comparison, in fall 2019 average math scores in our sample were above or on the typical growth line. This, in combination with results from the second research question, suggests that the rate of math growth for fall 2020 was reduced relative to typical growth and more in-line with a reduced rate of growth consistent with the partial absenteeism projections.

We further translated the projections into changes in percentile rank relative to the typical learning (50th percentile) scenario and compared those estimates to observed percentile changes since fall 2019. Table 7 presents a comparison of the projected and observed percentile ranks drops for reading and math. Consistent with the figures, students' observed percentile drops fell above both projected scenarios in reading and fell closest to partial absenteeism in math.

5. Conclusion

The analyses described in this technical report and accompanying research brief are an initial snapshot of how the COVID-19 pandemic has impacted students' achievement and growth. We examined math and reading test scores from students in over 8,000 US public schools and found evidence of continued gains in reading and unfinished learning in math, with the greatest impacts observed in grades 3-5. While it is difficult to compare these findings to other initial studies[x,xi] due to differences in tests used and learning loss metrics reported, we see consistent evidence of sizeable impacts on math achievement.

There are several important limitations worth noting. Most importantly, we only observed and therefore only used fall 2020 achievement data for the schools that tested in both 2019-20 and fall 2020 and the students within those schools that tested in both years (for details on missing assessment data see our attrition analysis[vii]). While we have presented some initial bounds for our RQ1 results based on few plausible scenarios to fill in the missing fall 2020 test scores, further research is needed to better understand and quantify the impacts of student attrition on our findings.

Further, we had access to limited demographic information on students, and while we examine racial/ethnic differences in MAP Growth percentiles, we have not yet examined growth by these same groups. Finally, these results represent only short-term impacts on math and reading skills. Future work will be needed to examine longer-term academic impacts as well as to measure students' social and emotional learning as students continue to face unprecedented challenges due to the pandemic.

We plan to conduct ongoing research to understand the impacts of the pandemic on student growth and achievement as more data become available. Future research will examine differences in students' achievement across the 2020-21 school year disaggregated by whether schools were able to re-open in person or remotely at the start of the school year. An additional line of work will examine in further depth how middle school math was impacted by COVID disruptions and the specific skills that were most impacted. Through this ongoing work, we seek to provide data to inform evidence-based policies to support our students, teachers, and families on the path to recovery.

Table 1. Description of the First Analytic Sample for Reading (Used in Research Question 1)

Grade	Male	White	Black	Other Race	Hispanic	Asian	Sample Size		
							Students	Schools	Districts
Fall 2019 Sample									
3	0.51	0.51	0.16	0.12	0.18	0.04	404,894	5,533	2,018
4	0.51	0.51	0.16	0.11	0.19	0.04	399,434	5,366	2,014
5	0.51	0.50	0.16	0.11	0.19	0.04	409,111	5,105	1,993
6	0.51	0.52	0.15	0.11	0.17	0.04	371,423	2,996	1,881
7	0.51	0.52	0.15	0.11	0.18	0.04	357,059	2,584	1,764
8	0.51	0.52	0.15	0.11	0.18	0.03	348,203	2,567	1,741
Fall 2020 Sample									
3	0.51	0.52	0.15	0.11	0.18	0.04	348,934	5,533	2,018
4	0.51	0.51	0.15	0.11	0.19	0.04	344,263	5,366	2,014
5	0.51	0.51	0.15	0.11	0.19	0.04	347,432	5,105	1,993
6	0.51	0.53	0.14	0.11	0.17	0.04	303,544	2,996	1,881
7	0.51	0.53	0.14	0.11	0.18	0.04	295,120	2,584	1,764
8	0.51	0.53	0.14	0.11	0.18	0.04	297,235	2,567	1,741

Note: To be included in this analytic sample, we required that a school tested at least ten students in a given grade in both fall 2019 and fall 2020.

Table 2. Description of the First Analytic Sample for Math (Used in Research Question 1)

Grade	Male	White	Black	Other Race	Hispanic	Asian	Sample Size		
							Students	Schools	Districts
Fall 2019 Sample									
3	0.51	0.50	0.16	0.12	0.19	0.04	441,301	6,001	2,134
4	0.51	0.49	0.16	0.11	0.19	0.04	447,049	6,013	2,139
5	0.51	0.49	0.16	0.11	0.20	0.04	462,520	5,766	2,132
6	0.51	0.51	0.16	0.11	0.18	0.04	433,165	3,422	2,061
7	0.51	0.51	0.16	0.11	0.18	0.04	420,810	2,971	1,960
8	0.51	0.51	0.16	0.11	0.18	0.04	394,133	2,888	1,899
Fall 2020 Sample									
3	0.51	0.50	0.15	0.11	0.20	0.04	391,909	6,001	2,134
4	0.51	0.50	0.15	0.11	0.20	0.04	399,121	6,013	2,139
5	0.51	0.49	0.15	0.11	0.20	0.04	405,860	5,766	2,132
6	0.51	0.52	0.15	0.11	0.18	0.04	371,394	3,422	2,061
7	0.51	0.52	0.15	0.11	0.18	0.04	360,853	2,971	1,960
8	0.51	0.52	0.15	0.11	0.18	0.03	333,778	2,888	1,899

Note: To be included in this analytic sample, we required that a school tested at least ten students in a given grade in both fall 2019 and fall 2020.

Table 3. Description of the Second Analytic Sample (Used in Research Questions 2 and 3)

Grade (F19)	Grade (F20)	Male	White	Black	Other Race	Hispanic	Asian	Sample Size		
								Students	Schools	Districts
Reading Sample										
3	4	0.51	0.50	0.14	0.12	0.20	0.04	325,662	5,593	1,942
4	5	0.51	0.50	0.14	0.11	0.21	0.04	325,308	5,344	1,911
5	6	0.51	0.53	0.14	0.11	0.19	0.04	260,986	3,008	1,751
6	7	0.51	0.52	0.15	0.10	0.19	0.04	260,447	2,546	1,616
7	8	0.51	0.52	0.15	0.10	0.19	0.03	237,481	2,473	1,576
Math Sample										
3	4	0.51	0.51	0.15	0.12	0.19	0.04	329,752	5,669	1,961
4	5	0.51	0.51	0.15	0.11	0.20	0.04	325,346	5,364	1,938
5	6	0.51	0.53	0.14	0.11	0.18	0.04	257,667	3,019	1,762
6	7	0.51	0.53	0.14	0.10	0.18	0.04	260,857	2,525	1,619
7	8	0.51	0.53	0.14	0.10	0.19	0.04	258,290	2,466	1,570

Note: To be included in the second analytic sample, we required that a student was assessed in fall 2019, winter 2020, and fall 2020.

Table 4. School Characteristics of the Fall 2020 NWEA Sample of Schools and the U.S. Public Schools Serving Each Grade Level

Group	Grade	Number of schools	Average School Enrollment	% FRPL	% White	% Asian	% Hispanic	% Black	% LEP	% Gifted	% Disabled	City	Rural	Suburb	Town
Fall 2020 NWEA Sample	3	5,533	478	0.50	0.64	0.03	0.17	0.14	0.10	0.06	0.14	0.27	0.28	0.33	0.12
Fall 2020 NWEA Sample	4	5,366	479	0.50	0.64	0.03	0.18	0.13	0.10	0.06	0.14	0.26	0.29	0.32	0.12
Fall 2020 NWEA Sample	5	5,105	482	0.50	0.64	0.03	0.18	0.13	0.10	0.06	0.14	0.26	0.29	0.32	0.12
Fall 2020 NWEA Sample	6	2,996	511	0.48	0.69	0.03	0.14	0.12	0.06	0.08	0.15	0.20	0.39	0.27	0.14
Fall 2020 NWEA Sample	7	2,584	529	0.47	0.69	0.03	0.14	0.13	0.05	0.09	0.15	0.20	0.40	0.25	0.15
Fall 2020 NWEA Sample	8	2,567	529	0.47	0.68	0.02	0.14	0.13	0.05	0.09	0.15	0.20	0.40	0.24	0.15
U.S. public schools	3	57,859	449	0.55	0.55	0.04	0.22	0.17	0.12	0.04	0.15	0.30	0.27	0.33	0.11
U.S. public schools	4	57,962	448	0.55	0.55	0.04	0.22	0.17	0.11	0.04	0.15	0.30	0.27	0.32	0.11
U.S. public schools	5	57,894	449	0.55	0.54	0.04	0.22	0.17	0.11	0.05	0.15	0.30	0.27	0.32	0.11
U.S. public schools	6	44,151	456	0.56	0.55	0.04	0.21	0.18	0.09	0.05	0.17	0.29	0.31	0.28	0.12
U.S. public schools	7	35,949	452	0.56	0.55	0.03	0.20	0.19	0.07	0.06	0.18	0.29	0.34	0.26	0.12
U.S. public schools	8	36,329	462	0.56	0.55	0.03	0.20	0.19	0.07	0.06	0.18	0.29	0.34	0.25	0.12

Note: FRPL=free or reduced priced lunch, LEP=limited English proficiency. This table is presented for the schools included in the first analytic sample for reading, but results are highly similar for across the various analytic samples. The school characteristics were retrieved from a school-level covariate data file produced by the Stanford Education Data Archive (SEDA) version 3.0 (Reardon et al., 2019). The sources of the variables are the Common Core of Data (CCD) collected by the National Center for Educational Statistics and the U.S. Department of Education (ED) Civil Rights Data Collection (CRDC). The U.S. public school population comparison for each grade was determined by limiting to the schools that offered a given grade.

Table 5. Classification of Students' Math RIT Scores into Achievement Quintiles by Term for Students Moving from Grades 3-4

		Fall 2020 Quintile					
		1-20	21-40	41-60	61-80	81-99	Total
Winter 2020 Quintile	1-20	11.8%	2.2%	0.5%	0.2%	0.2%	**15.0%**
	21-40	5.6%	7.3%	2.7%	0.6%	0.3%	**16.4%**
	41-60	1.7%	7.3%	8.0%	2.9%	0.5%	**20.5%**
	61-80	0.4%	3.0%	9.2%	10.9%	2.5%	**25.9%**
	81-99	0.1%	0.3%	1.9%	8.3%	11.7%	**22.3%**
	TOTAL	**19.5%**	**20.0%**	**22.3%**	**23.0%**	**15.2%**	**100.0%**

Note: N=329,752. The green boxes display students who moved to a higher quintile in fall 2020 relative to their winter 2020 quintile ("gainers"), the blue boxes display students who stayed in the same quintile of the distribution between winter 2020 and fall 2020 ("maintainers"), and the red boxes display students that moved down a quintile in fall 2020 ("sliders"). In total, 12.7 percent of grade 4 students were gainers in math, 49.6 were maintainers, and 37.7 percent of students were sliders.

Table 6. Results from the Sensitivity Analysis for Research Question 1

Grade	Subject	Lower bound	Observed Estimate	Upper Bound
3	Reading	-2	0	1
4	Reading	-3	-2	0
5	Reading	-2	-1	0
6	Reading	-2	0	1
7	Reading	-1	1	2
8	Reading	-2	1	2
3	Math	-10	-9	-8
4	Math	-11	-10	-9
5	Math	-9	-9	-7
6	Math	-8	-6	-5
7	Math	-7	-5	-4
8	Math	-6	-6	-4

Note: Lower bound, observed estimate, and upper bound are sets of estimated differences in median percentile rank between students who tested in fall 2019 and their same-grade peers in fall 2020 (among schools that tested in both terms). Observed estimate was calculated using only observed test scores. Lower bound and upper bound estimates were calculated after imputing four sets of fall 2020 test scores for students who tested in fall 2019 but not in fall 2020 and are based on the largest and smallest difference, respectively, from the observed fall 2019 median percentile.

Table 7. Percentile Changes Based on NWEA Projections Compared to Observed Percentile Drops Within Longitudinal Sample

Subject	Fall 2019 Grade	Fall 2020 Grade	Projected Fall 2020 Percentile (Relative to Typical)		Fall 2020 Percentile Drop (Relative to Fall 2019)
			Partial Absenteeism	COVID Slide	
Reading	3	4	-2	-8	-1
Reading	4	5	-4	-7	-3
Reading	5	6	-5	-7	-2
Reading	6	7	-6	-7	-1
Reading	7	8	-6	-6	-2
Math	3	4	-4	-19	-9
Math	4	5	-6	-20	-11
Math	5	6	-5	-20	-11
Math	6	7	-7	-14	-4
Math	7	8	-6	-9	-4

Note: Observed percentile drops were calculated with the in-tact cohort of students followed from fall 2019 to fall 2020. The projected percentile drops were calculated as the difference between the typical learning projection in fall 2020 (50^{th} percentile) and the associated percentile with the projected fall 2020 RIT score under the partial absenteeism and COVID-19 slide scenarios.

Figure 1: MAP Growth Percentiles by Grade Level in Fall 2019 and Fall 2020

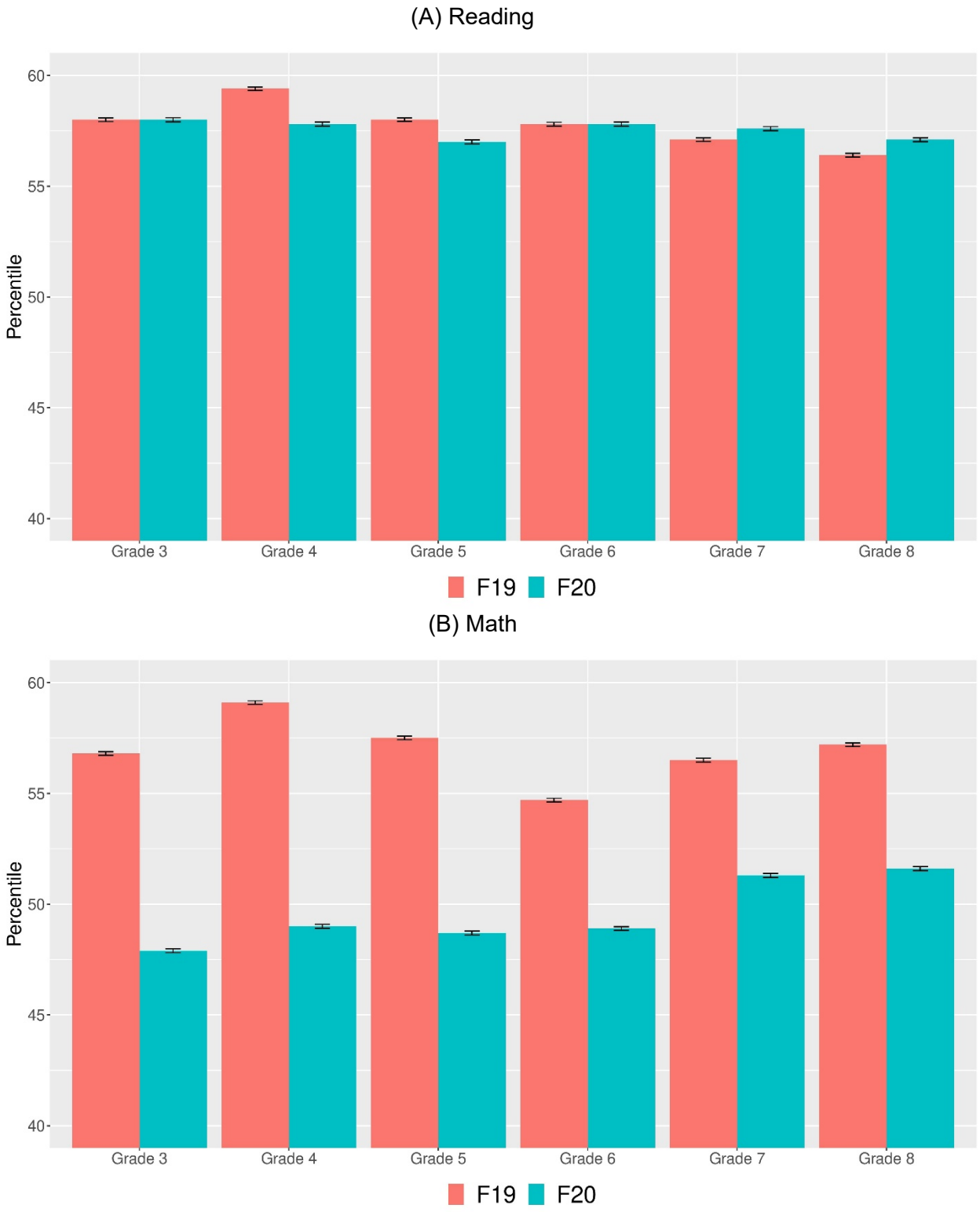


Figure 2: MAP Growth Percentiles by Grade Level and Race/Ethnicity in Fall 2019 and Fall 2020

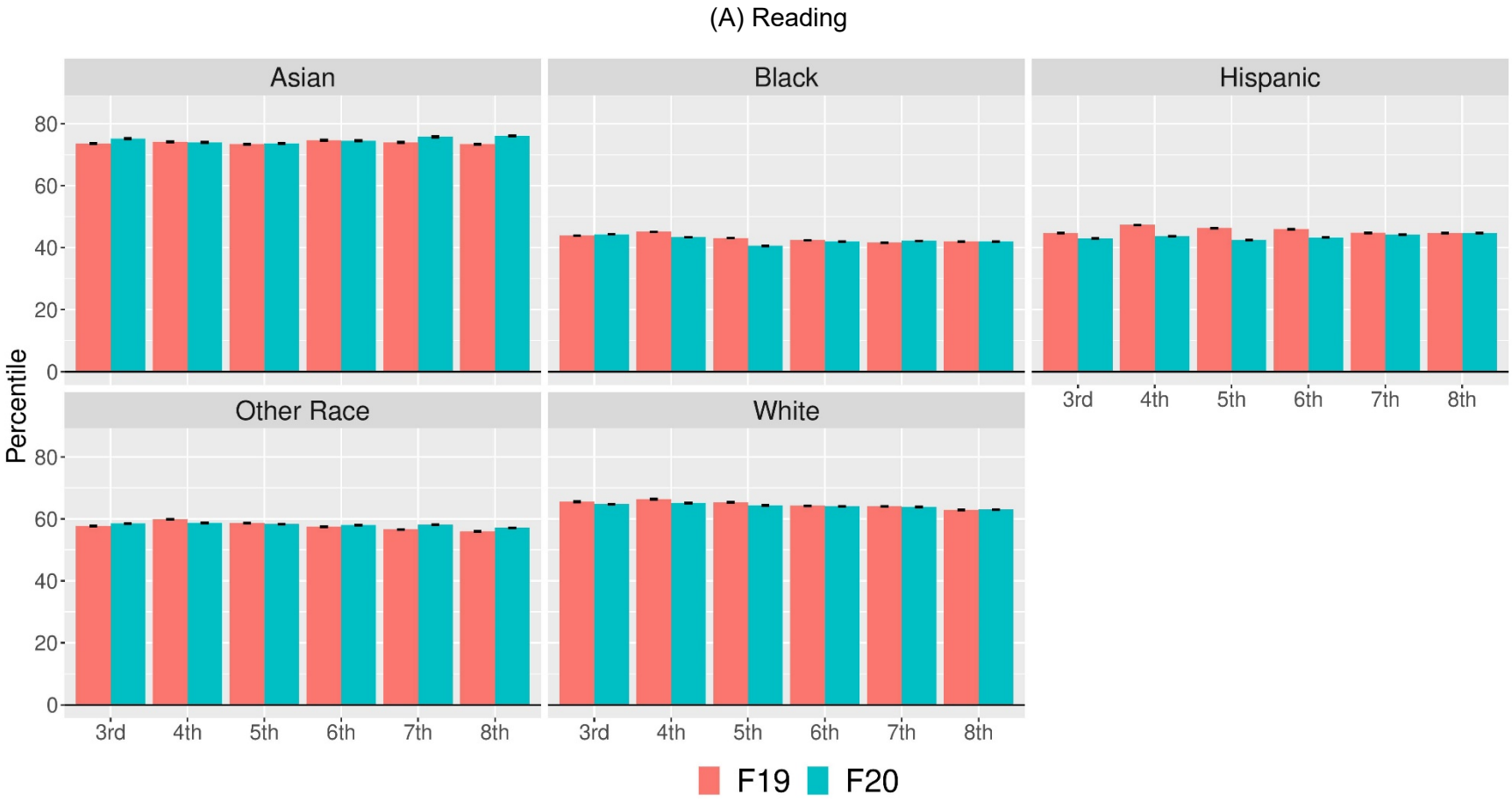


Figure 2: MAP Growth Percentiles by Grade Level and Race/Ethnicity in Fall 2019 and Fall 2020 (Continued)

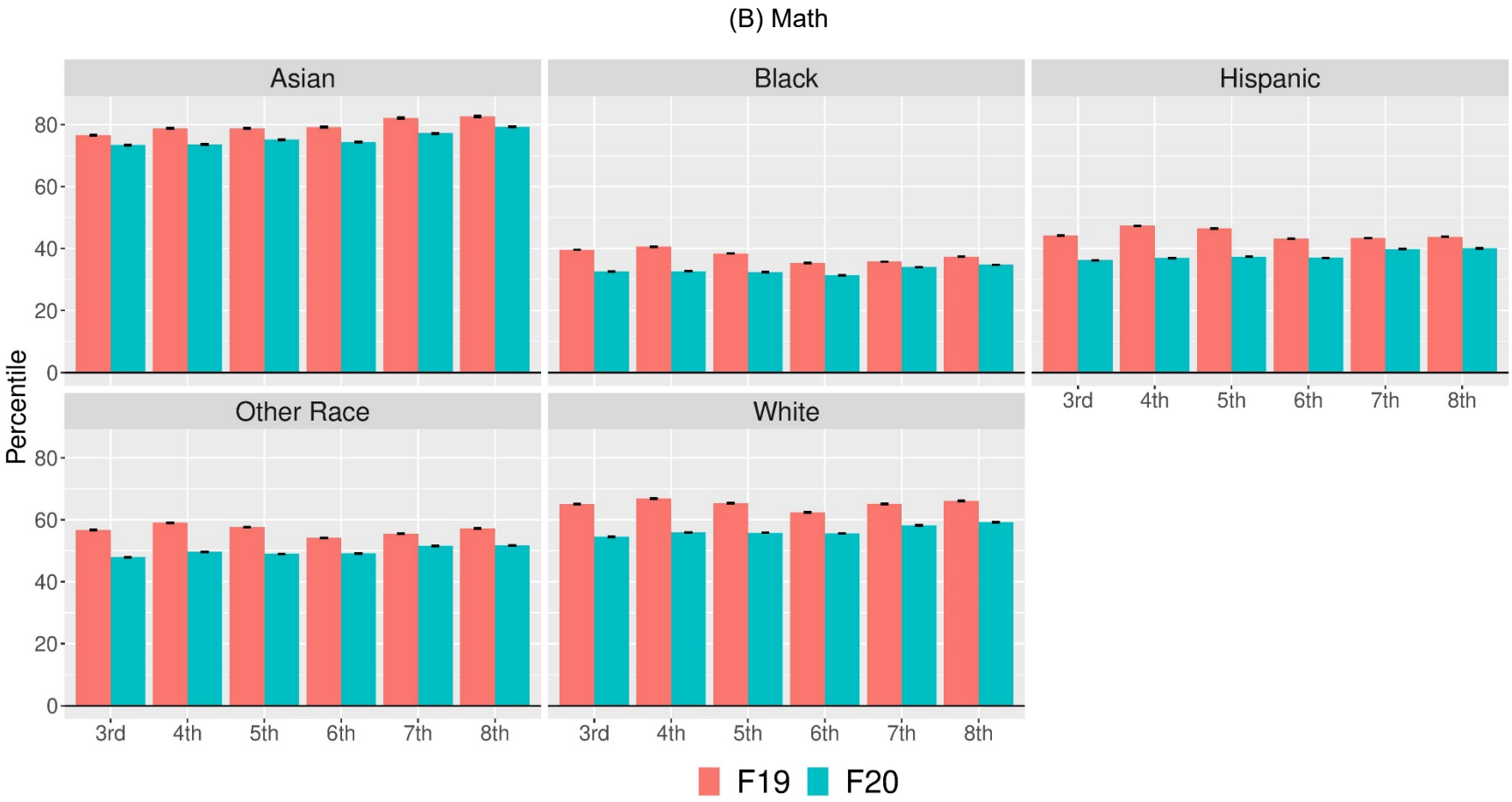


Figure 3: MAP Growth Percentiles by Grade Level and School Poverty Level in Fall 2019 and Fall 2020

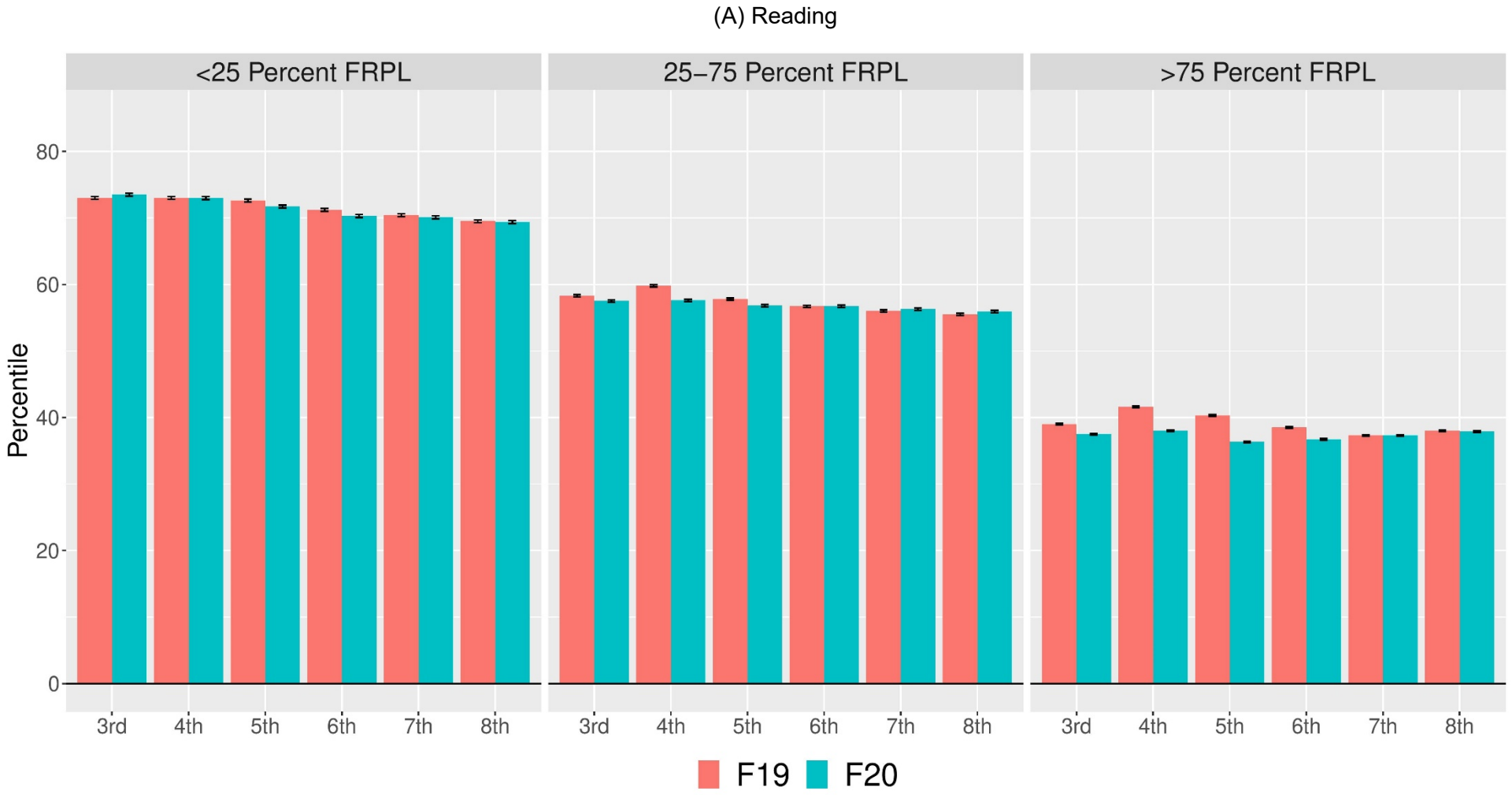


Figure 3: MAP Growth Percentiles by Grade Level and School Poverty Level in Fall 2019 and Fall 2020 (Continued)

(B) Math

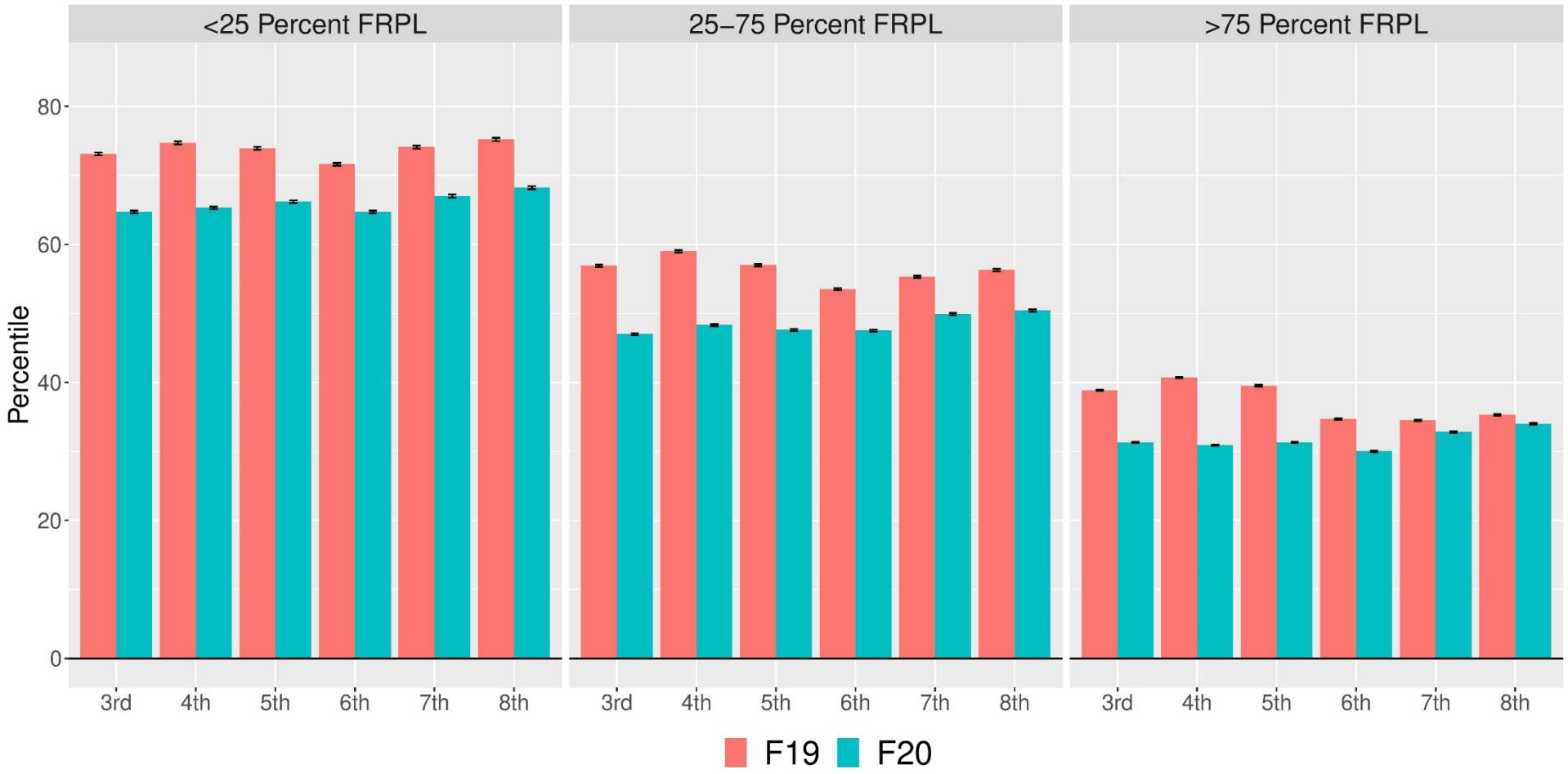


Figure 4: MAP Growth Percentiles by Grade Level and Urbanicity in Fall 2019 and Fall 2020

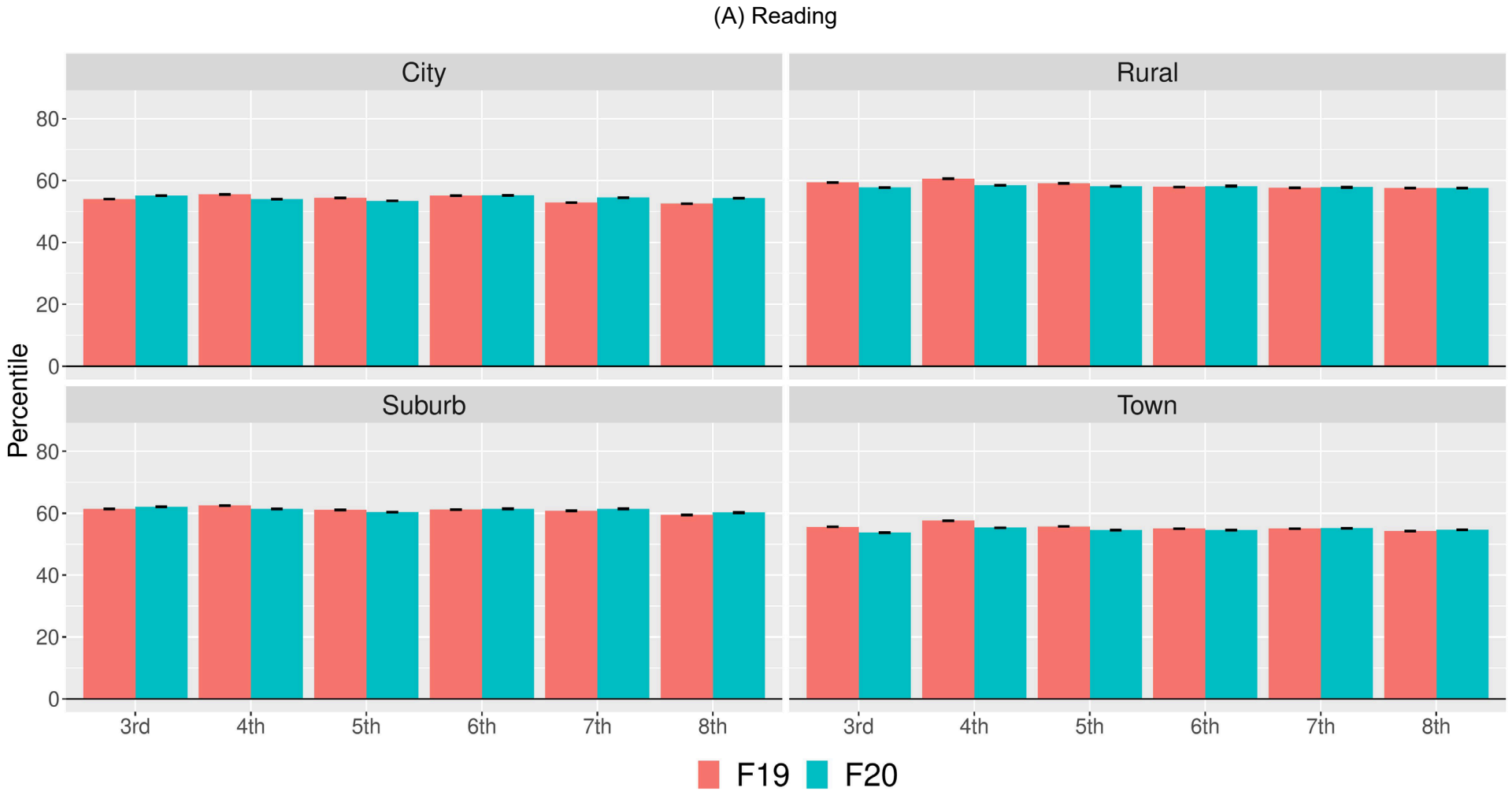


Figure 4: MAP Growth Percentiles by Grade Level and Urbanicity in Fall 2019 and Fall 2020 (Continued)

(B) Math

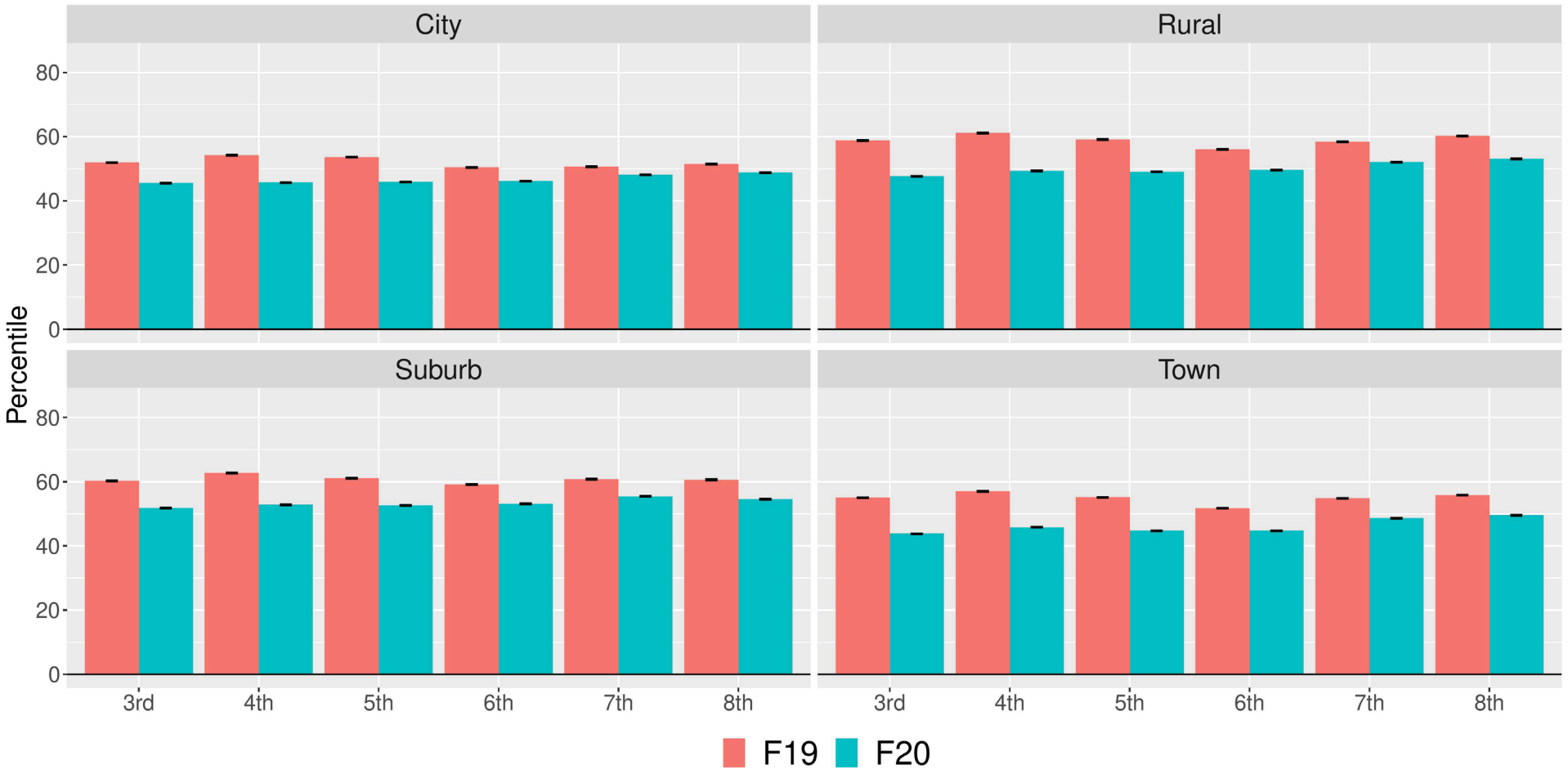
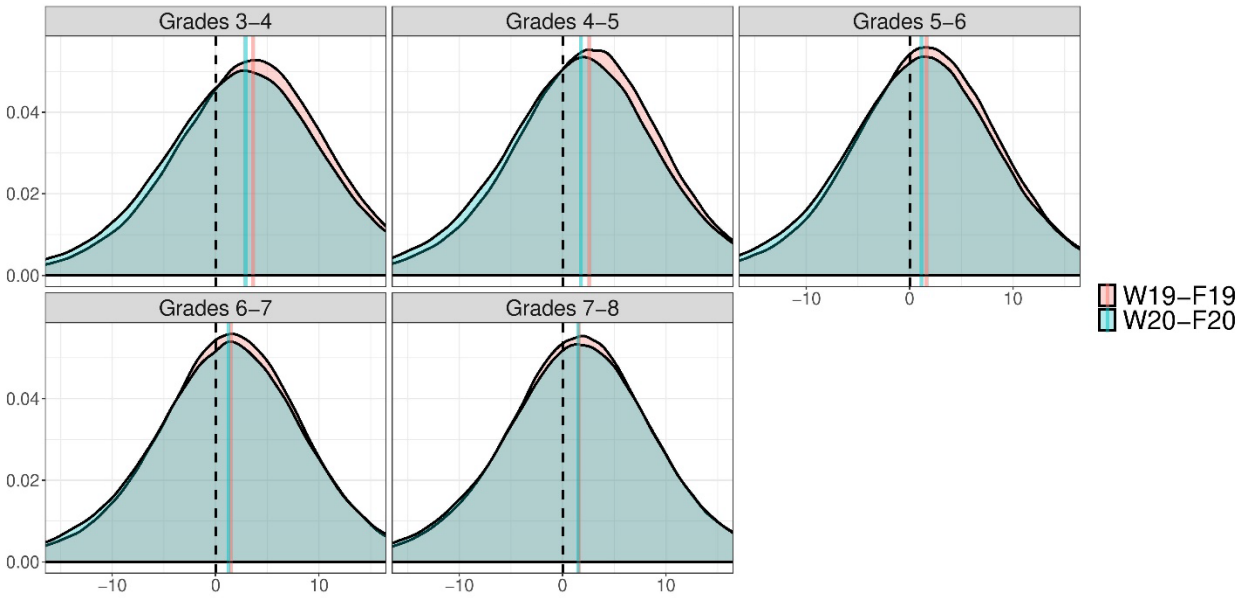


Figure 5: Distribution of Within-student Change from Winter 2019-Fall 2019 vs. Winter 2020-Fall 2020

(A) Reading



(B) Math

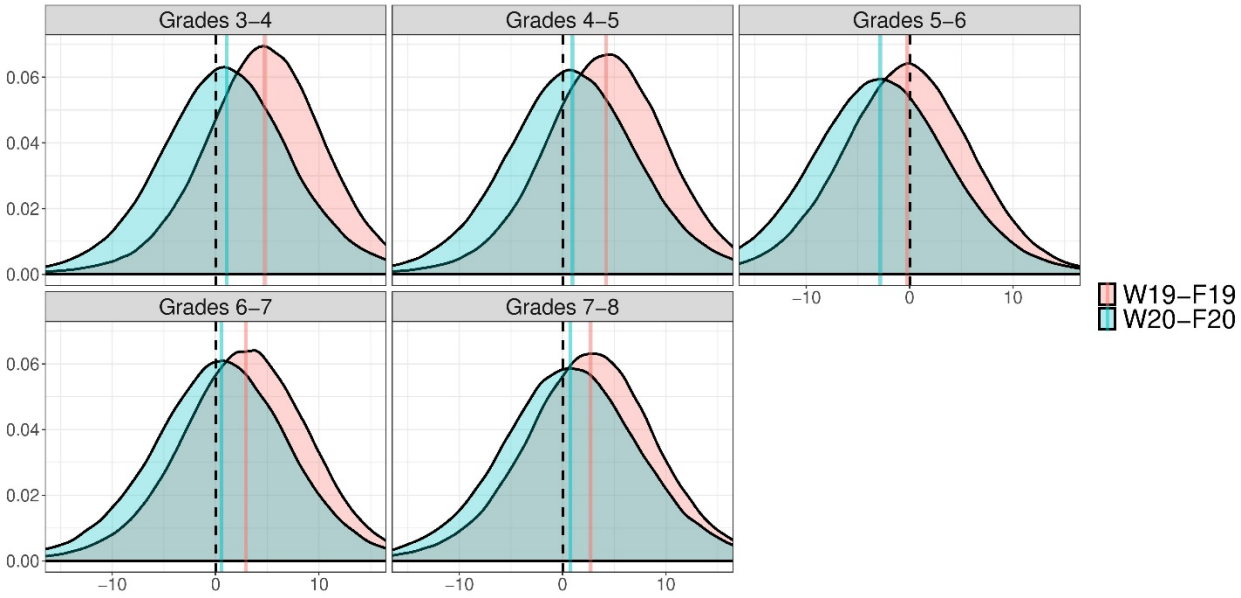
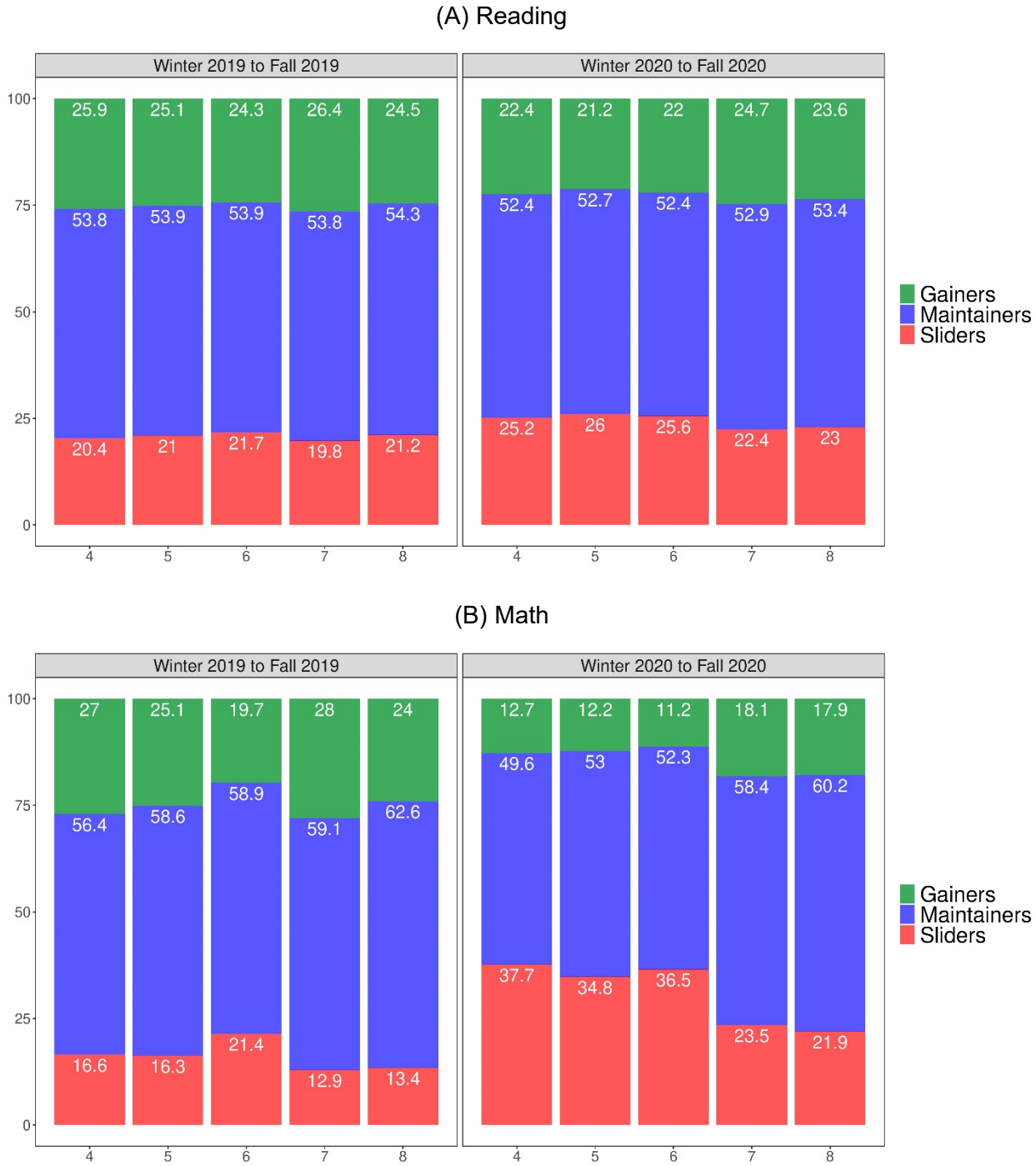
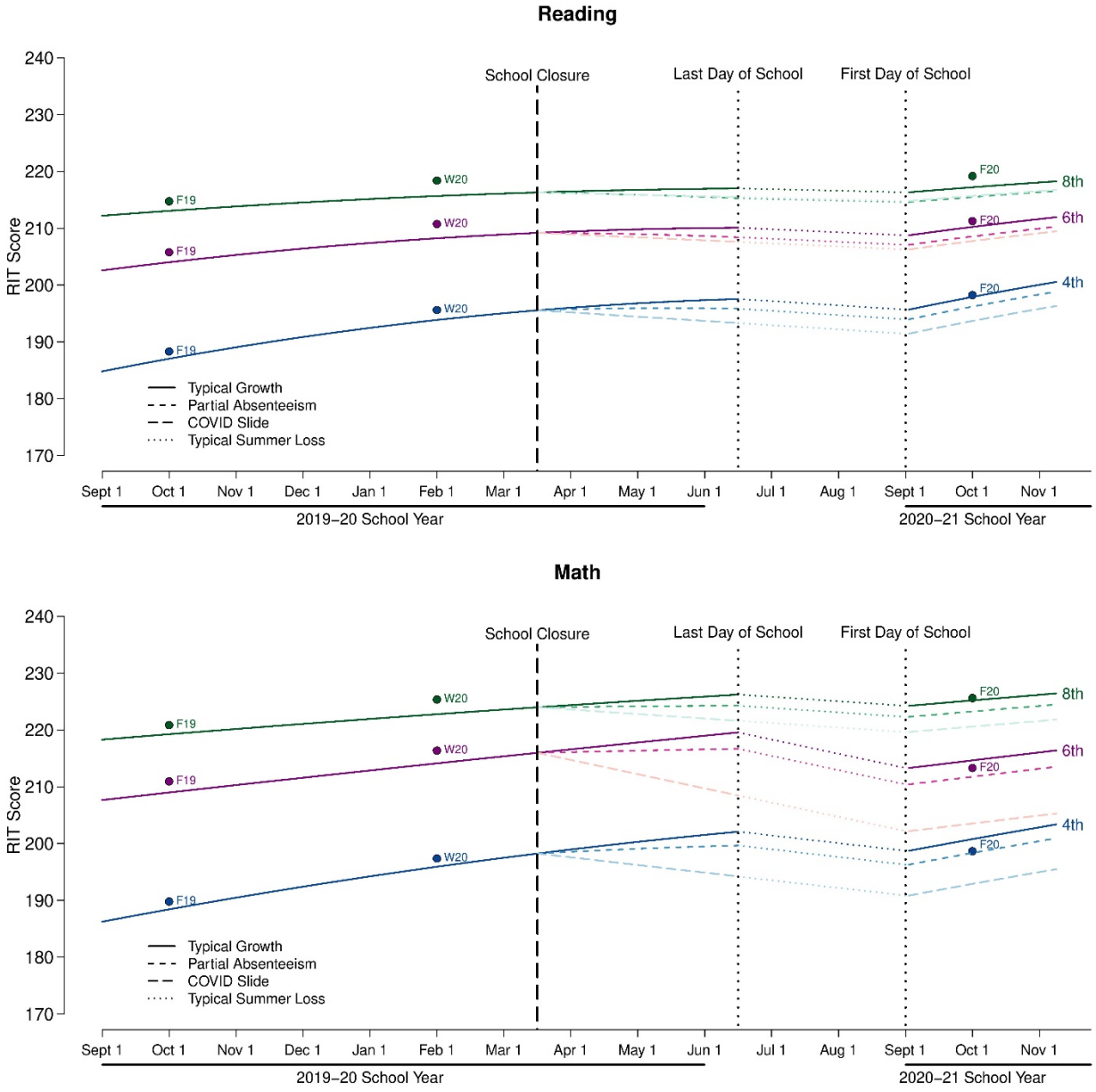


Figure 6: Percentage of Students by Grade Level Shifting Their Relative Position in the Test Score Distribution Comparing Winter 2019-Fall 2019 vs. Winter 2020-Fall 2020



Note: The x-axis displays the grade that students were in the fall (2019 or 2020) and the y-axis displays the percentage of students within each group (Gainers, Maintainers, Sliders).

Figure 7: Average Observed RIT Scores Mapped onto the NWEA COVID Slide Projections



Note: The projected lines were estimated based on student data collected in the 2017-18 and 2018-19 school years. The observed points (RIT scores averages from the testing windows with the most observed data) reflect data from our study sample collected during 2019-20 and 2020-21, which was higher performing than average relative to the overall NWEA sample. More information about each projected scenario is available in our paper.[ii] While one additional scenario ("full absenteeism") was considered in the source paper, we chose to omit it here to simplify the figure.

6. References

[i] Kuhfeld, M. & Tarasawa, B. (2020). The COVID-19 slide: What summer learning loss can tell us about the potential impact of school closures on student academic achievement. NWEA. https://www.nwea.org/content/uploads/2020/05/Collaborative-Brief_Covid19-Slide-APR20.pdf

[ii] Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549-565. https://doi.org/10.3102/0013189X20965918.

[iii] Estimates of learning loss in the 2019-2020 school year. (2020). The Center for Research on Education Outcomes. Stanford University. https://credo.stanford.edu/sites/g/files/sbiybj6481/f/short_brief_on_learning_loss_final_v.3.pdf

[iv] Dorn, E., Hancock, B, Sarakatsannis, J, & Viruleg, E. (2020). COVID-19 and student learning in the United States: The hurt could last a lifetime. https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-student-learning-in-the-united-states-the-hurt-could-last-a-lifetime#

[v] Kuhfeld, M., Tarasawa, B., Johnson, A., Ruzek, E., & Lewis, K. (2020). Learning during COVID-19: Initial findings on students' reading and math achievement and growth. NWEA. https://www.nwea.org/research/publication/learning-during-covid-19-initial-findings-on-students-reading-and-math-achievement-and-growth/

[vi] Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., Jang, H., Chavez, B., Buontempo, J., & DiSalvo, R. (2019). Stanford Education Data Archive (Version 3.0). http://purl.stanford.edu/db586ns4974.

[vii] Johnson, A. & Kuhfeld, M. (2020). Fall 2019 to fall 2020 MAP Growth attrition analysis. NWEA. https://www.nwea.org/research/publication/fall-2019-to-fall-2020-map-growth-attrition-analysis/

[viii] Kuhfeld, M., Lewis, K., Meyer, P., & Tarasawa, B. (2020). Comparability analysis of remote and in-person MAP Growth testing in fall 2020. NWEA. https://www.nwea.org/research/publication/comparability-analysis-of-remote-and-in-person-map-growth-testing-in-fall-2020/

[ix] Thum, Y. M., & Kuhfeld, M. (2020). NWEA 2020 MAP Growth achievement status and growth norms for students and schools. NWEA Research Report. Portland, OR: NWEA.

[x] Maldonado, J. & De Witte, K. (2020). The effect of school closures on standardised student test outcomes. https://lirias.kuleuven.be/3189074?limo=0

[xi] Huff (2020). National Data Quantifies Impact of COVID Learning Loss: Raises Questions about At-Home Testing. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/ca-impact-of-covid-learning-loss-fall-2020.pdf