IES Institute of Education Sciences

Estimating Changes to Student Learning in Illinois Following Extended School Building Closures due to the COVID-19 Pandemic

Regional Educational Laboratory Midwest

At American Institutes for Research

REL 2022–131
U.S. DEPARTMENT OF EDUCATION

A Publication of the National Center for Education Evaluation and Regional Assistance at IES

IES Institute of Education Sciences

Regional Educational Laboratory Midwest

At American Institutes for Research

Estimating changes to student learning in Illinois following extended school building closures due to the COVID-19 pandemic

Francie Streich, Jingtong Pan, Cong Ye, Jingyan Xia

December 2021

The COVID-19 pandemic disrupted the education of students in Illinois and around the nation. Leaders at the Illinois State Board of Education and in Illinois public school districts want to better understand how student learning changed during the pandemic. This study examines data from 17 Illinois districts over five years, including four years prior to the pandemic, to measure how student learning changed in fall 2020 relative to fall terms prior to the pandemic. The study demonstrates how learning changed in both mathematics and reading for students in grades 3–8, as well as how these changes varied across student characteristics and district size. The study found that students in grades 4–8 scored lower than expected in mathematics following the onset of the pandemic, after adjusting for other factors. The magnitude varied by grade level. Larger estimated changes in learning occurred in grades 6–8 than in grades 4 and 5. Students in grades 3–8 did not experience any statistically significant changes in learning in reading. A further analysis of learning in mathematics showed that changes in learning varied across students with different characteristics but were unrelated to district size. The study findings should be interpreted with caution, especially when generalizing to the population of Illinois districts and students. The study includes a small number of districts, and the students in these districts differ from the statewide population of students.

Why this study?

As a result of the COVID-19 pandemic, all public and nonpublic school buildings in Illinois were closed by executive order on March 17, 2020.[1] Illinois districts rapidly shifted to distance education for the remainder of the 2019/20 school year. Many districts continued in this format at the start of 2020/21. Research is needed to understand how the shift to distance education—alongside other challenges faced by students during the pandemic—affected student learning. Prior research has studied the effect of fully online distance education on K–12 student outcomes. These studies found mixed results (Ahn & McEachin, 2017; Chingos & Schwerdt, 2014; Hart et al., 2019; Heissel, 2016; Heppen et al., 2011, 2017; Woodworth et al., 2015) and may not be applicable in the current context. In addition to the change in instructional format, the pandemic resulted in many families experiencing job loss (Kochhar, 2020), food insecurity (Bitler et al., 2020), and psychological distress related to fear and isolation caused by the pandemic (Johns Hopkins University Bloomberg School of Public Health, 2020)—all of which may affect student learning (Picou & Marshall, 2007; Shores & Steinberg, 2017; Stevens & Schaller, 2011; Winicki & Jemison, 2003).

The Illinois State Board of Education requested this study to better understand how student learning changed following the onset of the pandemic and how changes in student learning may have varied by the characteristics of students and school districts in Illinois. The findings focus on changes in student learning during the early months of the pandemic, from the onset of the pandemic in March 2020 to the beginning of the 2020/21 school

[1] The executive order was issued by Illinois Governor J. B. Pritzker on March 13, 2020, and went into effect on March 17, 2020. A second executive order, issued on April 17, 2020, closed school buildings for the remainder of the 2019/20 school year.

year. State education leaders and district and school administrators can use these findings to inform decisions about how to target resources and to plan supports for student groups that experienced the largest reductions in learning. Furthermore, this study contributes to a growing evidence base about changes in student learning following the onset of the COVID-19 pandemic.

Research questions

The study is based on data from 17 districts in Illinois, which together serve about 3 percent of students in the state. The study explored two research questions related to changes in learning among students in grades 3–8 in these districts during the pandemic:

1. How had student learning in mathematics and reading changed as of fall 2020 following extended school building closures due to the COVID-19 pandemic?
2. Did these changes vary by student characteristics and district size?

Definitions of key terms are in box 1. The data sources, sample, and methods used to answer the research questions are in box 2.

Box 1. Key terms

Changes in student learning. Differences between students' expected test scores (based on how similar students scored in earlier years) and students' actual test scores on standardized interim assessments.

COVID-19 era. The period following school building closures in March 2020 through the time when students took the NWEA MAP assessment in fall 2020 (typically in September).

Northwest Evaluation Association Measures of Academic Progress (NWEA MAP). A standardized interim assessment that measures achievement in mathematics and reading (Northwest Evaluation Association, 2009). The assessment is adaptive and nationally normed. Districts commonly administer this assessment in the fall, winter, and spring to students in grades 3–8.

Pre-COVID-19. The period before school buildings closed in March 2020 in response to the pandemic.

Statistically significant. Indicates when a finding has a high probability of being real and not occurring by chance.

Unfinished learning. Refers to when a group of students' actual test scores on standardized interim assessments were lower than expected based on how similar students scored in earlier years.

Box 2. Data sources, sample, and methods

Data sources. The study used a combination of fall Northwest Evaluation Association Measures of Academic Progress (NWEA MAP) data provided by the participating districts, and student, school, and district records provided by the Illinois State Board of Education (ISBE). A full list of data elements, data sources, and variables used in the study is in appendix A.

ISBE provided the following data:

- Test score records from 2015/16 to 2020/21, including fall NWEA MAP mathematics and reading scores. These data were provided to ISBE by the participating districts.
- Data on the characteristics of students who were in grades 3–8 between 2015/16 and 2020/21, including school identification numbers, student identification numbers, and their demographic characteristics (age, gender, grade level, race/ethnicity, eligibility for the national school lunch program, special education status, and English learner status). ISBE merged these student characteristics with the test score records provided by the participating districts and masked student identification numbers before transferring data to Regional Educational Laboratory Midwest.
- Data on the characteristics of schools, including school identification numbers, district identification numbers, the percentage of students from different racial/ethnic groups, the percentage of students eligible for the national school lunch program, and the percentage of students who are English learner students.
- Data on the characteristics of districts, including district identification number, the percentage of students from different racial/ethnic groups, the percentage of students eligible for the national school lunch program, district size, and region.[1]

Sample. A nonrandom sample of 17 districts provided data for students in grades 4–8; the study sample represents 2 percent of the 850 school districts in Illinois that serve students in at least one of these grades. In 2019/20 these 17 districts served 3.3 percent of students in Illinois public schools. Fourteen of the 17 districts also provided data for students in grade 3; the study sample represents 1.6 percent of the 850 school districts in Illinois that serve grade 3. In 2019/20, these 14 districts served 2.4 percent of students in Illinois public schools. Each grade-level analysis includes students in that grade at different points in time, and each grade-level sample includes five cohorts of students in that grade. In total, the study team used data from the 2015/16–2020/21 school years, including observations of 148,392 students for analyses of mathematics and 148,498 students for analyses of reading. The study team excluded students with missing data. All analyses were conducted separately for each grade level. For additional details about the analytic sample, see appendix A.

Methodology. To answer research question 1 about changes in student learning, the study used regression models to estimate the relationship between the onset of the COVID-19 pandemic and NWEA MAP scores. The regression models in this study adjust for the relationships between NWEA MAP scores and student background characteristics (age, gender, race/ethnicity, eligibility for the national school lunch program, special education status, English learner status, and prior-year fall test score), school characteristics (the percentage of students who are White, the percentage of students eligible for the national school lunch program, and the percentage of students who are English learner students), district characteristics (the percentage of students who are White, the percentage of students eligible for the national school lunch program, and district size), and a time trend. The study team converted regression model estimates to days of instruction for ease of interpretation (see appendix A). The study team tested the sensitivity of the findings for research question 1 to the use of two different weighting approaches (see appendix C).

To answer research question 2 about the variation in changes in student learning by student and district characteristics, the study used regression models similar to those used for research question 1 but allowed the relationship between student learning and the COVID-19 pandemic to vary across multiple student characteristics and district size. The figures in this report illustrate whether there are statistically significant associations between these characteristics and changes in learning. Positive (or negative) associations do not imply that students with a particular characteristic scored higher (or lower) than a student without that characteristic. Rather, associations indicate that the change in learning for students with a particular characteristic, relative to what was expected of them, was better (either more positive or less negative) or worse (either less positive or more negative) than the change in learning for students without that characteristic, relative to what was expected of them.

Generalizability. This study is based on data from a small sample of districts in Illinois. The study districts, and the students they serve, differ from the population of districts in Illinois in multiple ways. None of the districts in the study sample is in the southern regions of the state, and Chicago Public Schools is not in the study sample. Chicago Public Schools enrolls 18 percent of all public school students in Illinois, and these students are different, on average, from those of the rest of the state. The

districts in the study sample have larger percentages of students eligible for the national school lunch program, smaller percentages of White students, larger percentages of Hispanic students, and larger percentages of English learner students than the statewide population. As such, the study findings are not generalizable to the whole state, but rather to districts that serve student populations similar to the sample of students in this study. More details on the characteristics of the districts and students in the study sample are in appendixes A and C.

Other limitations. The reader should keep in mind three additional limitations when considering study findings. First, the study does not provide causal estimates of the impact of the COVID-19 pandemic on student learning. All students in Illinois experienced the effects of the pandemic at the same time, meaning that there is no comparison group against which to compare academic outcomes. Instead, the study uses data from students in several years prior to the pandemic as a comparison group and adjusts for student, school, and district characteristics as well as a time trend to provide estimates of the relationship between the pandemic and student learning. Learning changes following the onset of the pandemic may have been related to unmeasured factors rather than to the pandemic. Second, changes in the mode of administration for the fall NWEA MAP assessment may have affected the findings. Before the pandemic, the NWEA MAP was administered to all students in person. In fall 2020, many districts administered the assessment to students remotely. The study data do not identify the administration format used for each student, and the study cannot control for or isolate findings based on mode of administration.[2] Third, the study sample only includes four pre-COVID-19 time points, which may result in an inaccurate estimate of the time trend leading to the expected COVID-19–era test scores. For instance, if student test scores tend to fluctuate from year to year, then the study may not estimate changes in learning precisely. Readers should not focus on a specific estimate of changes in instructional days of learning but rather on the direction of and patterns in estimated learning changes.

Notes

1. Illinois is divided into 10 geographic regions (see map A1 in appendix A).

2. NWEA studied the differences between in-school and remote testing and found that the different testing formats produced largely comparable scores (Meyer, 2020). However, the Regional Educational Laboratory Midwest study team was unable to perform the same comparison to determine whether scores would have been comparable in the study sample.

Findings

This section presents the main findings from the study. Caution should be used when generalizing the findings to the population of grade 3–8 students in Illinois because the findings are based on a small sample of Illinois districts whose students are different from the statewide population of students. Additional findings and supporting analyses are in appendixes B and C.

In fall 2020, following the onset of the COVID-19 pandemic, students scored lower than expected in mathematics after adjusting for other factors, but not in reading

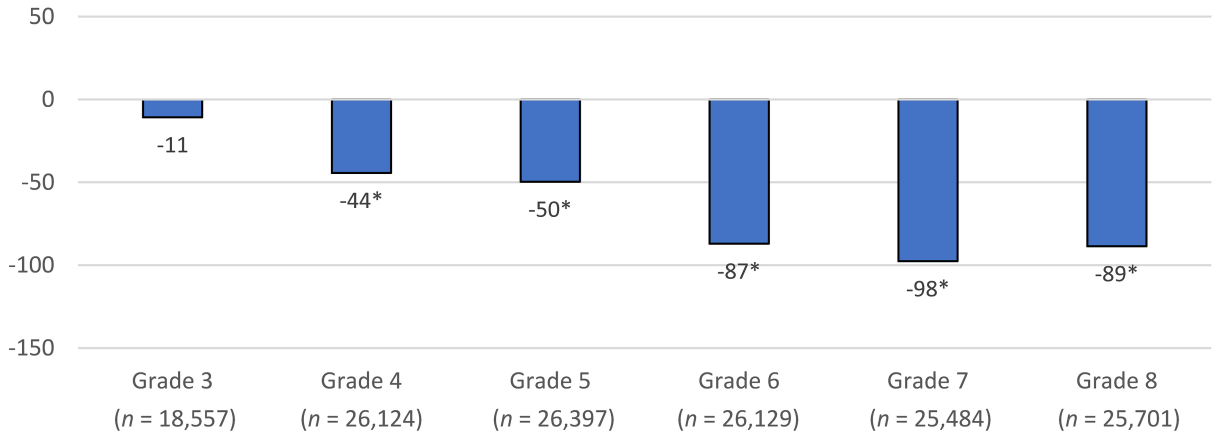
In fall 2020, after approximately six months of disruption to school and life due to the COVID-19 pandemic, student performance on the Northwest Evaluation Association Measures of Academic Progress (NWEA MAP) mathematics assessment was lower than expected in grades 4–8, adjusting for other factors (figure 1). In other words, actual test scores for students in these grade levels were lower than expected based on how similar students scored in earlier years, indicating that there was unfinished learning. The estimated amount of unfinished learning was largest in the middle grades (grades 6–8). In grade 6, students experienced unfinished learning equivalent to 87 fewer days of instruction, or 49 percent of a school year. In grade 7, students experienced unfinished learning equivalent to 98 fewer days of instruction, or 56 percent of a school year. Finally, in grade 8, students experienced unfinished learning equivalent to 89 fewer days of instruction, or 51 percent of a school year.

In reading, the pattern was different. Student scores were not statistically significantly different from expected in grades 3–8, after adjusting for other factors (figure 2). Although the estimated changes in learning for grades 3, 6, and 8 were large (equivalent to 39 more days of instruction in grade 3 and 34 and 59 fewer days of instruction in grades 6 and 8, respectively), they were not statistically significant. This is because the variability in reading scores

among students in these grades in fall 2020 was sufficiently large that the study team cannot determine, with certainty, that the change in learning is different from zero.

Figure 1. Students in grades 4–8 scored lower than expected in mathematics, after adjusting for other factors, fall 2020

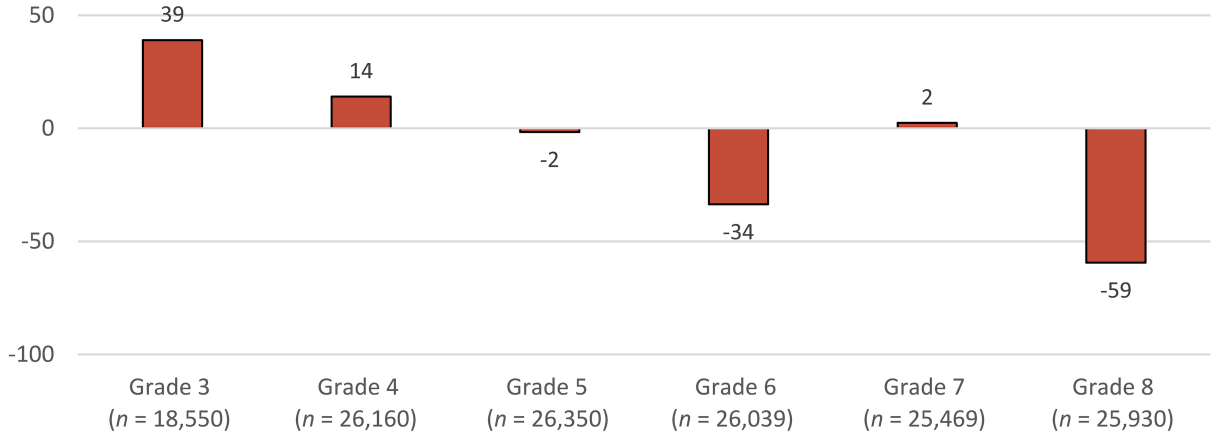
Estimated change in learning
(Days of instruction)



* Estimated change in learning was significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure 2. Students in grades 3–8 did not score differently from expected in reading, after adjusting for other factors, fall 2020

Estimated change in learning
(Days of instruction)



Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. None of these estimates is statistically significant at $p < .05$.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Because the sample of students in these analyses was different from the population of students in grades 3–8 in Illinois in several ways, the study team tested the sensitivity of the study's findings to the inclusion of weights that were developed to improve the representativeness of the sample. The results of the weighted and unweighted analyses are similar (see figures C5 and C6 in appendix C). Like the unweighted analyses, students in grades 4–8 scored lower than expected in mathematics, and students in grades 3–8 did not score differently than expected

in reading. The study team also tested the sensitivity of the study's findings to its primary approach for addressing missing information in the student-level data within the participating districts. To do so, the team included weights that adjusted for missing student-level data but that did not attempt to improve the representativeness of the sample for the statewide population of students. The results of this weighting approach are largely similar to those in figures 1 and 2 (see figures C7 and C8).[2]

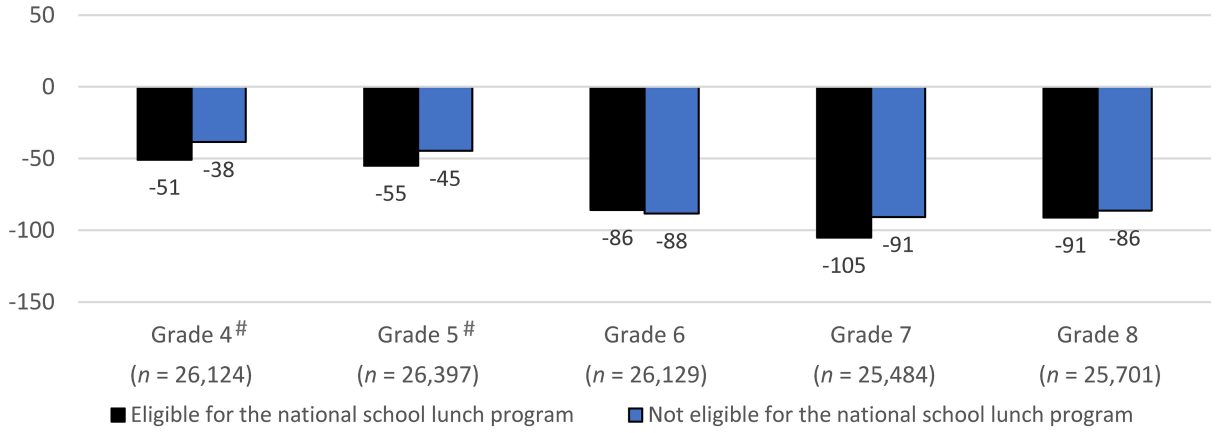
Student characteristics, such as English learner status and special education status, are associated with changes in learning in mathematics, but the direction of these relationships varied by grade level

The following figures (figures 3–6) illustrate comparisons between student groups where the study found statistically significant reductions in learning in mathematics in grades 4–8. The narrative discusses all differences between student groups that are statistically significant, but focuses on those that are largest in magnitude. All the results discussed in the narrative are statistically significant, unless otherwise stated. Given interest among state education leaders and district and school administrators in allocating resources and planning supports for students who experienced the largest amounts of unfinished learning, these comparisons can help guide resources to students most in need. Results of similar comparisons for learning in mathematics in grade 3 and learning in reading in grades 3–8 are shown in appendix B.

In grades 4 and 5, students who were eligible for the national school lunch program experienced more unfinished learning in mathematics than students who were not eligible, after adjusting for other factors. In grades 4–8, students experienced unfinished learning in mathematics regardless of their eligibility for the national school lunch program, after adjusting for other factors (figure 3). However, in grades 4 and 5, students who were eligible for the national school lunch program experienced more unfinished learning than students who were not eligible.

Figure 3. In grades 4 and 5, students who were eligible for the national school lunch program experienced more unfinished learning in mathematics than students who were not eligible, after adjusting for other factors, fall 2020

Estimated change in learning
(Days of instruction)



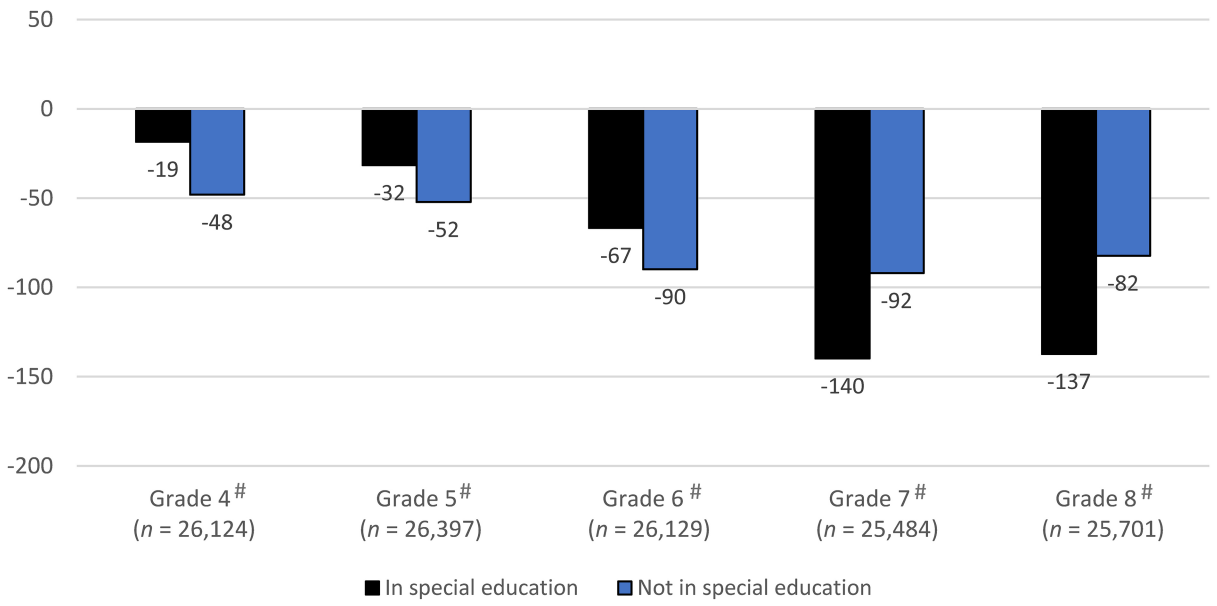
Differences between student groups were statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. All student groups experienced statistically significant changes in learning at $p < .05$.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

[2] Applying the student-level weights resulted in a statistically significant reduction in learning in mathematics in grade 3 (see figure C7 in appendix C) and a statistically significant reduction in learning in reading in grade 6 (see figure C8). In both cases, the main findings were not statistically significant.

In grades 7 and 8, students in special education had more unfinished learning in mathematics than students who were not in special education, after adjusting for other factors. In each grade level, students' special education status was associated with changes in learning in mathematics (figure 4). However, the direction of the relationship varied by grade level. Students in special education in lower grades (grades 4–6) had less unfinished learning than students who were not in special education in these grades. However, students in special education in higher grades (grades 7 and 8) had more unfinished learning than students who were not in special education in these grades.

Figure 4. In grades 7 and 8, students in special education had more unfinished learning in mathematics than students who were not in special education, after adjusting for other factors, fall 2020

Estimated change in learning
(Days of instruction)

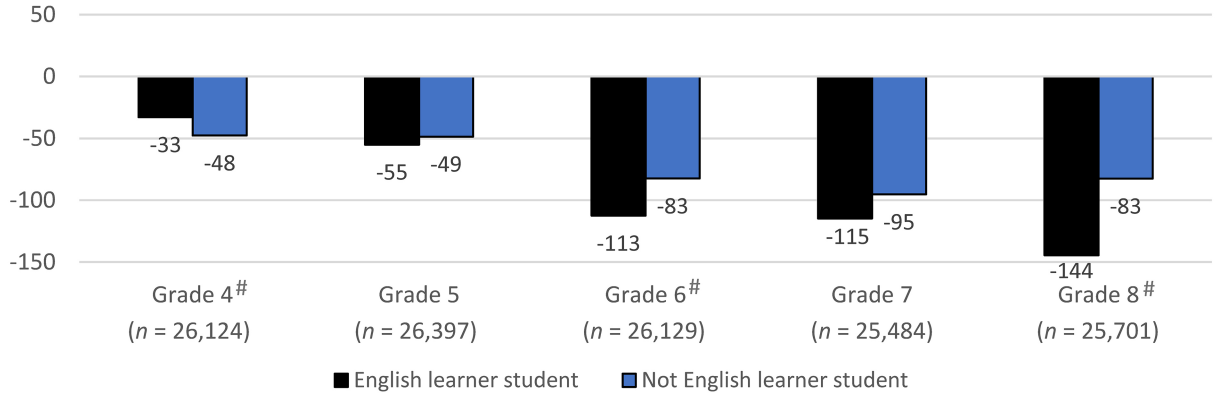


Differences between student groups were statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. All student groups experienced statistically significant changes in learning at $p < .05$.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

In grades 6 and 8, English learner students had more unfinished learning in mathematics than students who were not English learner students, after adjusting for other factors. In grades 4, 6, and 8, students' English learner status was associated with changes in learning in mathematics, after adjusting for other factors (figure 5). Like the findings for students in special education, the direction of this relationship differed between students in lower and higher grades. In grade 4, English learner students had less unfinished learning in mathematics than their peers who were not English learner students. However, in grades 6 and 8, English learner students had more unfinished learning in mathematics than their peers who were not English learner students.

Figure 5. In grades 6 and 8, English learner students had more unfinished learning in mathematics than their peers who were not English learner students, after adjusting for other factors, fall 2020

Estimated change in learning
(Days of instruction)

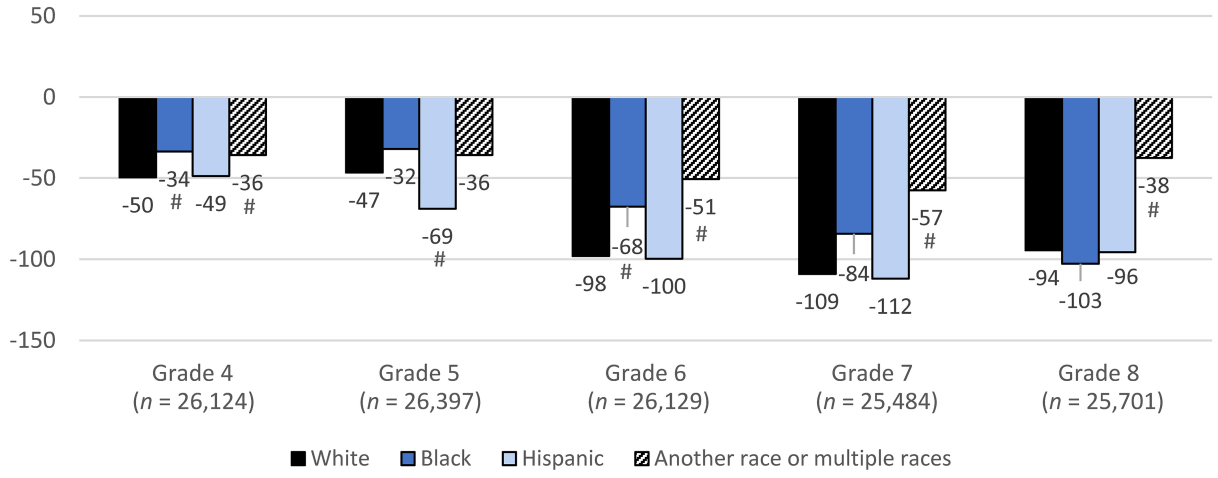


Differences between student groups were statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. All student groups experienced statistically significant changes in learning at $p < .05$.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

In grades 4 and 6–8, White students had more unfinished learning in mathematics than students of another race or multiple races, after adjusting for other factors. In grades 4 and 6–8, White students had more unfinished learning in mathematics than students of another race or multiple races (figure 6). The latter group includes students who are American Indian, Asian, Pacific Islander, or two or more races. White students in these grade levels had unfinished learning equivalent to 50–109 fewer days instruction, whereas estimated changes in learning for students of another race or multiple races were not significantly different from zero. The study also found that White students experienced more unfinished learning than Black students in grades 4 and 6 and less unfinished learning than Hispanic students in grade 5.[3]

Figure 6. In grades 4 and 6–8, White students had more unfinished learning in mathematics than students of another race or multiple races, after adjusting for other factors, fall 2020

Estimated change in learning
(Days of instruction)



The difference between racial/ethnic group and White students was statistically significant at $p < .05$.

Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. Another race or multiple races includes students who are American Indian, Asian, Pacific Islander, or two or more races. All student groups experienced statistically significant changes in learning at $p < .05$ except for students of another race or multiple races in grades 4–8 and Black students in grades 4 and 5.

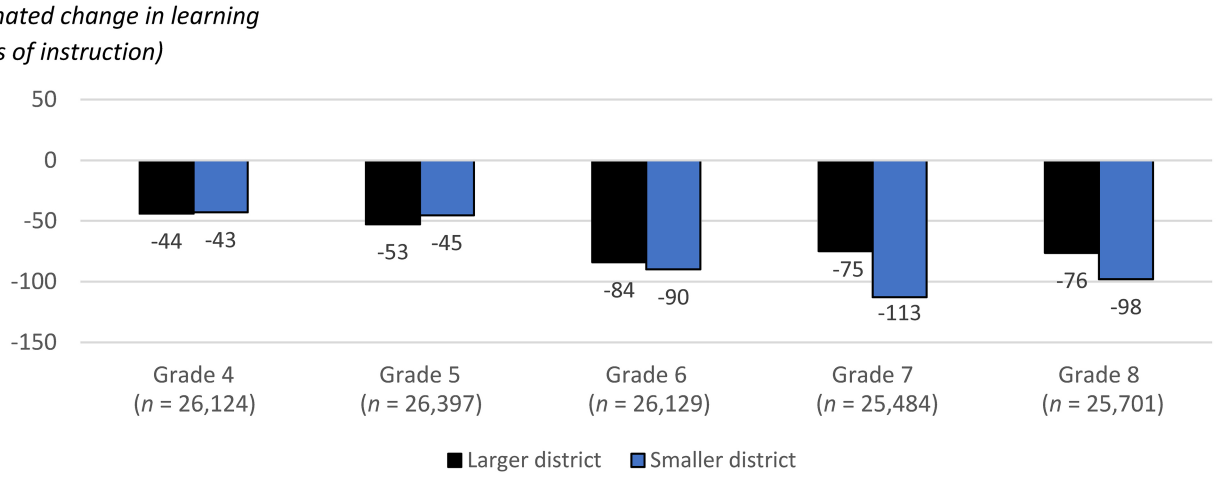
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

[3] The study team conducted a sensitivity analysis in which it only controlled for student-level race/ethnicity and a time trend. The study team found that none of the differences between racial/ethnic groups were significant. The relationships between student race/ethnicity and changes in learning, which are illustrated in figure 6, only hold after accounting for other factors, including eligibility for the national school lunch program, English learner student status, special education status, and prior academic achievement.

District size was not related to changes in learning in mathematics, after adjusting for other factors

The study compared changes in student learning between smaller and larger districts. Larger districts are defined as those with total enrollment greater than or equal to the average district-level enrollment of approximately 4,000 students. District size was not related to changes in learning in mathematics in any grades studied (figure 7).

Figure 7. District size was not related to changes in learning in mathematics in any grade, after adjusting for other factors, fall 2020



Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. Larger districts are those with total enrollment greater than or equal to the average district-level total enrollment; smaller districts are smaller than the average. The average district total enrollment is 3,998. All student groups experienced statistically significant changes in learning at $p < .05$. None of the differences between larger and smaller districts is statistically significant at $p < .05$.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Implications

The study findings have two main implications for Illinois education leaders as they allocate resources and plan supports for their students while continuing to navigate the COVID-19 pandemic. The first implication is that school and district leaders may want to identify and implement evidence-based strategies that can accelerate learning in mathematics. In five of the six grade levels analyzed, students scored worse than expected in mathematics following the onset of the pandemic, after adjusting for other factors. Other recent research similarly found declines in mathematics performance following the onset of the pandemic (Kuhfeld et al., 2020; Renaissance Learning, 2020). More specifically, school and district leaders may want to focus their resources on strategies that can accelerate mathematics learning for students with the greatest unfinished learning, including English learner students and students in special education in higher grades.

The second implication is that additional research is needed to further inform districts' and schools' efforts, as well as state-level efforts, to support students through the duration of the pandemic and beyond. Three kinds of additional research would be useful. First, the present study shows whether, and how much, student learning changed, but additional research is needed to understand why. As district and school administrators allocate resources and make plans to support students with learning recovery, they also may want to gather additional information—such as through surveys of or interviews with parents and students—that can help identify why student learning changed as it did. Administrators may want to focus their efforts on student groups that experienced the most unfinished learning. For example, the study found that students in higher grades had more unfinished learning in mathematics than those in lower grades, and that unfinished learning was especially large in higher grades for English learner students and students in special education. District and school administrators may want to focus on identifying the drivers of unfinished learning for these students, then subsequently on

identifying and providing appropriate assistance to support their learning recovery. Second, additional research is needed to understand changes in student learning for a more comprehensive and representative sample of students. The present study used data from a small sample of districts whose students differ from the Illinois population in important ways. Future research could explore opportunities for gathering and analyzing data from a larger sample of districts to produce findings that are more generalizable to the population of students in Illinois. The findings of this research could be more useful for planning that occurs at the state level. Third, additional research is needed to understand how student learning changed over a longer period of time. The present study focused on changes in student learning as of fall 2020, after approximately six months of disruption to school and life due to the COVID-19 pandemic. However, the pandemic has persisted well beyond this time. Many districts continued to serve students in remote or hybrid formats throughout the 2020/21 school year. Furthermore, many students and their families continued to experience other challenges associated with the pandemic, such as food insecurity and psychological distress. All of these could lead to changes in learning beyond those observed in this study.

References

Ahn, J., & McEachin, A. (2017). Student enrollment patterns and achievement in Ohio's online charter schools. *Educational Researcher, 46*(1), 44–57. https://eric.ed.gov/?id=EJ1132550

Bitler, M., Hoynes, H. W., & Schanzenbach, D. W. (2020). *The social safety net in the wake of COVID-19* (NBER Working Paper No. 27796). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w27796/w27796.pdf

Chingos, M. M., & Schwerdt, G. (2014). *Virtual schooling and student learning: Evidence from the Florida Virtual School* (PEPG Working Paper No. 14-02). Harvard Kennedy School. https://www.hks.harvard.edu/sites/default/files/Taubman/PEPG/research/PEPG14_02.pdf

Hart, C. M. D., Berger, D., Jacob, B., Loeb, S., & Hill, M. (2019). Online learning, offline outcomes: Online course taking and high school student performance. *AERA Open, 5*(1), 1–17. https://eric.ed.gov/?id=EJ1210509

Heissel, J. (2016). The relative benefits of live versus online delivery: Evidence from virtual algebra 1 in North Carolina. *Economics of Education Review, 53,* 99–115.

Heppen, J. B., Sorensen, N., Allensworth, E., Walters, K., Rickles, J., Taylor, S. S., & Michelman, V. (2017). The struggle to pass algebra: Online vs. face-to-face credit recovery for at-risk urban students. *Journal of Research on Educational Effectiveness, 10*(2), 272–296. https://eric.ed.gov/?id=ED604473

Heppen, J. B., Walters, K., Clements, M., Faria, A., Tobey, C., Sorensen, N., & Culp, K. (2011). *Access to algebra 1: The effects of online mathematics for grade 8 students* (NCEE 2012-4021). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://eric.ed.gov/?id=ED527394

Johns Hopkins University Bloomberg School of Public Health. (2020, June 3). Survey finds large increase in psychological distress reported among US adults during the COVID-19 pandemic. *ScienceDaily*. http://www.sciencedaily.com/releases/2020/06/200603132550.htm

Kochhar, R. (2020, June 9). *Hispanic women, immigrants, young adults, those with less education hit hardest by COVID-19 job losses*. Pew Research Center. https://www.pewresearch.org/fact-tank/2020/06/09/hispanic-women-immigrants-young-adults-those-with-less-education-hit-hardest-by-covid-19-job-losses/

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/content/uploads/2020/11/Collaborative-brief-Learning-during-COVID-19.NOV2020.pdf

Meyer, P. (2020). *Comparisons between remote testing and in-school testing for MAP Growth: A summary of results for spring 2020.* NWEA. https://www.nwea.org/resource-library/map-growth-research-guidance/comparisons-between-remote-testing-and-in-school-testing-for-map-growth-3

Northwest Evaluation Association. (2009). *Technical manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades.*

Northwest Evaluation Association. (2020). *2020 NWEA MAP Growth normative data overview.* https://teach.mapnwea.org/impl/MAPGrowthNormativeDataOverview.pdf

Picou, J. S., & Marshall, B. K. (2007). Social impacts of Hurricane Katrina on displaced K–12 students and educational institutions in coastal Alabama counties: Some preliminary observations. *Sociological Spectrum*, *27*(6), 767–780.

Renaissance Learning. (2020). *How kids are performing: Tracking the impact of COVID-19 on reading and mathematics achievement.* https://www.renaissance.com/how-kids-are-performing/

Shores, K., & Steinberg, M. P. (2017). *The impact of the Great Recession on student achievement: Evidence from population data* (CEPA Working Paper No. 17-09). Stanford Center for Education Policy Analysis. https://cepa.stanford.edu/sites/default/files/wp17-09-v201708.pdf

Stevens, A. H., & Schaller, J. (2011). Short-run effects of parental job loss on children's academic achievement. *Economics of Education Review*, *30*(2), 289–299. https://eric.ed.gov/?id=EJ916914

Winicki, J., & Jemison, K. (2003). Food insecurity and hunger in the kindergarten classroom: Its effect on learning and growth. *Contemporary Economic Policy*, *21*(2), 145–157.

Woodworth, J. L., Raymond, M. E., Chirbas, K., Gonzalez, M., Negassi, Y., Snow, W., & Van Donge, C. (2015). *Online charter school study.* Center for Research on Education Outcomes. https://credo.stanford.edu/sites/g/files/sbiybj6481/f/online_charter_study_final.pdf

REL 2022–131

December 2021

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-17-C-0004 by the Regional Educational Laboratory Midwest administered by the American Institutes for Research. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

IES Institute of Education Sciences

APPENDIXES

Regional Educational Laboratory Midwest

At American Institutes for Research

Estimating changes to student learning in Illinois following extended school building closures due to the COVID-19 pandemic

Appendix A. Data and methods

Appendix A provides additional details about the data and analysis methods used in the study.

Data sources

The study used multiple data sources.

Student-by-term-level Northwest Evaluation Association Measures of Academic Progress data. A nonrandom sample of 14 Illinois districts provided Northwest Evaluation Association Measures of Academic Progress (NWEA MAP) fall mathematics and reading assessment scores for all students in their districts in grades 2–8 between 2015/16 and 2020/21. The data from grade 2 and from 2015/16 were requested to include students' prior academic achievement as a covariate in the analyses. An additional three districts provided NWEA MAP scores for students in grades 3–8. These three districts are only included in analyses of grades 4–8 because the prior academic achievement covariate is not available for grade 3 students. The data also included school identification number, student identification number, name, date of birth, and grade level to allow the Illinois State Board of Education (ISBE) to merge the scores with other student demographic characteristics.

Student-by-year-level demographic data. ISBE provided demographic data for all students enrolled in the 17 participating districts in grades 3–8 between 2016/17 and 2020/21. These data included student age, grade level, gender, race/ethnicity, eligibility for the national school lunch program, English learner status, and special education status. These data included school identification numbers to facilitate merging with the school- and district-level data also provided by ISBE.

School- and district-by-year-level data. ISBE provided school and district characteristics for all Illinois districts serving students in grades 3–8 between 2016/17 and 2020/21. These school characteristics included the percentage of students who were White, the percentage of students eligible for the national school lunch program, and the percentage of students who were English learner students. The district characteristics included the percentage of students who were White, the percentage of students eligible for the national school lunch program, and district size.

Data preparation

ISBE merged student-by-term-level NWEA MAP data files with the student-by-year-level demographic data based on unique student IDs. In some cases, student IDs in the district files contained errors and could not be used. In these instances, ISBE merged the NWEA MAP records with the demographic records, when possible, using students' names and dates of birth. ISBE then replaced the student IDs with a masked student ID and removed name and date of birth before providing the combined student-level data file to Regional Educational Laboratory Midwest.

To prepare the data file for analysis, the study team took the following steps:

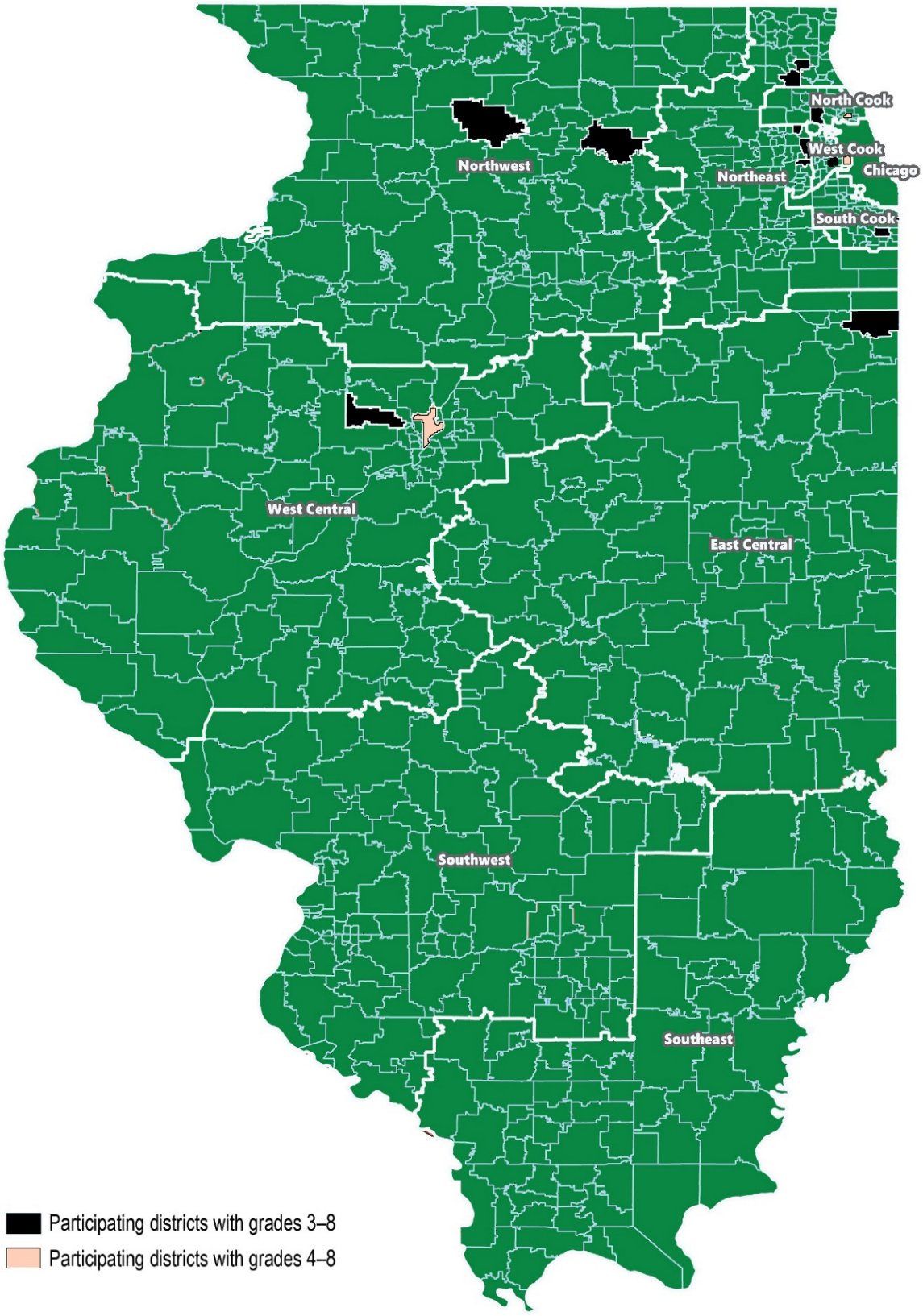
1. When students had more than one assessment score in the same term and same subject, the study team randomly selected one of the scores for inclusion in the analytic sample.
2. When students repeated a grade, the study team kept the earliest record of the student enrolled in a given grade level. For example, if a student enrolled in grade 4 in 2016/17 and again in 2017/18, the study team kept only the 2016/17 record for the student.
3. The study team excluded students who did not have an assessment score in either the fall term of a given school year or the fall term of the prior school year, because the analytic models adjust for prior academic achievement. For example, students were included in the analytic sample for the analysis of grade 4 reading scores only if they had a valid reading score in the fall of grade 4 and the fall of grade 3.
4. In three districts, most grade 3 students did not have assessment scores in grade 2. The study team removed these three districts from the grade 3 analytic sample.
5. The study team merged the student-level data with the school-level data and district-level data using the school and district indicators provided by ISBE.
6. For three of the 17 participating districts, one of the pre-COVID-19 time points is excluded from the analytic sample. In these districts, the average scores in each grade level in these time points were over one standard deviation (about 15 points, which is greater than full year of progress) higher or lower than the average scores from other time points for the same grade level. The study team was concerned that these districts had inadvertently swapped their students' mathematics and reading scores in these time points.[4] Because of this issue, the study team excluded fall 2016 data for two districts and fall 2018 data for one district.

Study sample

The study sample comprised students in grades 4–8 in 17 districts between 2016/17 and 2020/21. Among them, 14 districts also include students in grade 3. The 14 and 17 districts represent 1.6 percent and 2 percent, respectively, of the 850 districts in Illinois that serve students in grades 3–8. A map of Illinois and its regions with the location of these districts is in map A1.

[4] The differences would all be less than one standard deviation if those mathematics and reading results were switched.

Map A1. Illinois districts in study sample



Source: Authors' compilation.

The sample districts and their individual and aggregate characteristics are in table A1. The sample includes districts of varying sizes (ranging from 494 to 12,864 total students) and with varying student populations (ranging, for example, from 1 percent to 99 percent of students eligible for the national school lunch program). Altogether, the students in the 14 districts used in the grade 3 analyses and in the 17 districts used in the grades 4–8 analyses differ from the statewide population of students. For example, both samples of districts have smaller percentages of White students and larger percentages of English learner students than the state overall.

Table A1. Characteristics of Illinois districts in study sample

		Percentage of students who were:						
District	Size	Eligible for the national school lunch program	English learner students	In special education	White	Black	Hispanic	Another or multiple races
Brookwood SD 167	1,167	56	11	10	4	75	19	2
Butler SD 53	494	1	5	14	42	1	3	54
CCSD 62[a]	4,447	45	34	17	39	4	42	15
Cicero SD 99	10,689	92	55	14	1	2	96	1
DeKalb CUSD 428	6,604	60	15	16	40	23	29	8
Elmhurst SD 205	8,563	15	10	14	70	2	17	11
Elmwood CUSD 322	670	23	—	13	93	1	2	4
Hawthorn CCSD 73	3,937	26	23	15	43	3	26	28
Hoover-Schrum Memorial SD 157	858	99	14	15	1	75	21	3
Kildeer Countryside CCSD 96	3,273	5	18	10	49	2	9	40
Maywood-Melrose Park-Broadview 89	4,690	70	31	11	2	30	65	3
Momence CUSD 1	1,041	69	10	13	53	14	28	5
Oregon CUSD 220	1,493	45	1	15	84	1	10	5
Peoria SD 150[a]	12,864	79	7	17	20	57	12	11
Riverside SD 96	1,722	21	5	13	61	3	31	5
Skokie SD 69[a]	1,754	50	32	12	35	11	17	37
Wood Dale SD 7	1,058	56	41	11	39	3	53	5
Total (14 districts in grade 3 sample)	46,259	51	25	14	35	11	43	11
Total (17 districts in grades 4–8 samples)	65,324	56	22	14	32	20	36	12
Statewide total	1,957,018	49	12	15	48	17	27	8

— is not available. CCSD is consolidated community school district. CUSD is community unit school district. SD is school district.

Note: All characteristics were measured in 2019/20.

a. District not included in grade 3 analytic sample.

Source: Authors' compilation of data from the Illinois State Board of Education.

Missing data. Within the sample districts, some students were missing data and were excluded from the analysis. Missing data were noted only for fall NWEA MAP scores but not for student demographic characteristics or school- and district-level characteristics. The study team assessed the percentage of students enrolled in the sample districts in each grade level and year that were excluded from the mathematics and reading analyses due to missing data (tables A2 and A3, respectively). For both mathematics and reading, each grade level excludes between 14 percent and 17 percent of enrolled students due to missing data—either missing the score used as an outcome or the prior-year score used as a covariate. The study team excluded students who were missing data and conducted a complete case analysis. The study team conducted a sensitivity analysis for the approach to dealing with missing student-level data within the participating districts by using a response propensity weighting approach. The results were similar to the unweighted results presented in the main findings with all differences less than 3 estimated instructional days (see figures C7 and C8 in appendix C).

Table A2. Number of valid cases and percentage missing fall Northwest Evaluation Association Measures of Academic Progress mathematics scores

Grade level and year	Number of valid cases	Percentage missing outcome score	Percentage missing prior-year test score	Percentage missing any test score	Number of cases excluded due to missing data	Total enrollment
Grade 3						
2016/17	3,971	6	15	15	724	4,695
2017/18	3,922	6	15	16	721	4,643
2018/19	3,801	5	14	15	657	4,458
2019/20	3,632	7	15	17	728	4,360
2020/21	3,231	10	11	18	701	3,932
Total	18,557	7	14	16	3,531	22,088
Grade 4						
2016/17	5,349	6	14	15	974	6,323
2017/18	5,544	5	13	14	919	6,463
2018/19	5,459	6	13	14	891	6,350
2019/20	5,192	6	13	14	855	6,047
2020/21	4,580	13	11	20	1163	5,743
Total	26,124	7	13	16	4,802	30,926
Grade 5						
2016/17	5,276	6	13	14	861	6,137
2017/18	5,534	5	13	14	888	6,422
2018/19	5,479	6	12	14	858	6,337
2019/20	5,429	5	12	13	795	6,224
2020/21	4,679	12	10	18	1,020	5,699
Total	26,397	7	12	14	4,422	30,819
Grade 6						
2016/17	5,091	6	14	15	888	5,979
2017/18	5,324	6	13	14	874	6,198
2018/19	5,356	9	12	16	1,037	6,393
2019/20	5,447	5	12	13	825	6,272
2020/21	4,911	13	10	19	1,167	6,078
Total	26,129	8	12	15	4,791	30,920
Grade 7						
2016/17	4,825	7	17	19	1,153	5,978
2017/18	5,323	6	12	14	834	6,157
2018/19	5,111	9	15	19	1,180	6,291
2019/20	5,336	6	15	16	1,045	6,381
2020/21	4,889	14	10	19	1,162	6,051
Total	25,484	8	14	17	5,374	30,858
Grade 8						
2016/17	4,762	10	15	20	1,161	5,923
2017/18	5,335	6	13	14	903	6,238
2018/19	5,244	8	14	17	1,101	6,345
2019/20	5,295	6	14	15	964	6,259
2020/21	5,065	13	9	18	1,113	6,178
Total	25,701	9	13	17	5,242	30,943

Note: Grade 3 includes data from 14 districts. Grades 4–8 include data from an additional three districts, for a total of 17 districts.
Source: Authors' analysis of data provided by participating districts and the Illinois State Board of Education.

Table A3. Number of valid cases and percentage missing fall Northwest Evaluation Association Measures of Academic Progress reading scores

Grade level and year	Number of valid cases	Percentage missing outcome score	Percentage missing prior-year test score	Percentage missing any test score	Number of cases excluded due to missing data	Total enrollment
Grade 3						
2016/17	3,861	6	17	18	834	4,695
2017/18	3,897	6	15	16	746	4,643
2018/19	3,802	6	14	15	656	4,458
2019/20	3,639	6	15	17	721	4,360
2020/21	3,351	8	10	15	581	3,932
Total	18,550	6	14	16	3,538	22,088
Grade 4						
2016/17	5,322	6	15	16	1,001	6,323
2017/18	5,547	5	13	14	916	6,463
2018/19	5,451	6	13	14	899	6,350
2019/20	5,186	6	13	14	861	6,047
2020/21	4,654	12	11	19	1,089	5,743
Total	26,160	7	13	15	4,766	30,926
Grade 5						
2016/17	5,264	6	13	14	873	6,137
2017/18	5,535	6	13	14	887	6,422
2018/19	5,480	6	12	14	857	6,337
2019/20	5,422	5	12	13	802	6,224
2020/21	4,649	12	10	18	1,050	5,699
Total	26,350	7	12	15	4,469	30,819
Grade 6						
2016/17	5,097	6	13	15	882	5,979
2017/18	5,329	6	13	14	869	6,198
2018/19	5,296	9	12	17	1,097	6,393
2019/20	5,441	5	12	13	831	6,272
2020/21	4,876	14	10	20	1,202	6,078
Total	26,039	8	12	16	4,881	30,920
Grade 7						
2016/17	4,892	6	17	18	1,086	5,978
2017/18	5,317	6	13	14	840	6,157
2018/19	5,057	10	15	20	1,234	6,291
2019/20	5,277	6	16	17	1,104	6,381
2020/21	4,926	13	10	19	1,125	6,051
Total	25,469	8	14	17	5,389	30,858
Grade 8						
2016/17	5,033	5	14	15	890	5,923
2017/18	5,418	5	12	13	820	6,238
2018/19	5,152	10	14	19	1,193	6,345
2019/20	5,246	6	15	16	1,013	6,259
2020/21	5,081	13	9	18	1,097	6,178
Total	25,930	8	13	16	5,013	30,943

Note: Grade 3 includes data from 14 districts. Grades 4–8 include data from an additional three districts, for a total of 17 districts.
Source: Authors' analysis of data provided by participating districts and the Illinois State Board of Education.

Altogether, the analytic sample of students differs from the statewide population of students in multiple ways. For example, in 2018/19, the sample percentage of students who are White is lower than the population by 15 to 18 percentage points depending on the grade level. The sample percentage of students who are Hispanic is higher than the population by 14 to 22 percentage points depending on the grade level (table A4). In grades 3 and 6, the sample percentage of students who are English learner students is higher than the population by 6 percentage points. The analytic sample is similar to the statewide population of students in multiple ways. For example, the sample percentage of students who are eligible for the national school lunch program is within 5 percentage points of the population percentage in each grade level. The sample percentage of students who are receiving special education is within 4 percentage points of the population in each grade level.

Table A4. Student characteristics in the analytic sample and statewide population, 2018/19

	Percentage of students who were:						
Grade level and sample	Eligible for the national school lunch program	English learner students	In special education	White	Black	Hispanic	Another or multiple races
Grade 3							
Sample	51	25[a]	13	32[a]	9[a]	47[a]	12
State population	52	19	15	47	18	26	9
Grade 4							
Sample	53	18	13	30[a]	18	41[a]	11
State population	52	19	15	47	17	27	9
Grade 5							
Sample	55	18	13	31[a]	15	42[a]	12
State population	52	15	15	47	17	27	9
Grade 6							
Sample	56	16[a]	11	30[a]	16	44[a]	10
State population	52	10	15	47	17	28	8
Grade 7							
Sample	45	12	11	32[a]	15	42[a]	11
State population	50	8	15	47	17	27	9
Grade 8							
Sample	45	12	11	32[a]	15	42[a]	11
State population	49	7	14	48	16	27	9

Note: The percentages of students with each characteristic are similar (differences of 1 percentage point or less) in both the mathematics and reading analytic samples, so only the mathematics analytic sample characteristics are presented here.
a. The difference between the sample and statewide percentage is greater than 5 percentage points.
Source: Authors' analysis of data provided by participating districts and the Illinois State Board of Education.

The study team also examined the extent to which there were patterns in the missing data, focusing especially on whether there were higher rates of missingness across student groups in the COVID-19 era (2020/21) in comparison to the year immediately prior (table A5). One pattern emerged across all grade levels. Black students had different rates of missingness between 2019/20 and 2020/21. In grade 3, Black students had a higher rate of missing data in the 2019/20 school year than in the 2020/21 school year. By contrast, Black students in grades 4–8 had a higher rate of missing data in the COVID-19 era than in the pre-COVID-19 year. Differences were also found for other student groups, but not consistently across all grade levels. In grade 3, English learner students had a higher rate of missing data in the 2020/21 school year than in the 2019/20 school year. In grades 4–6, non–English learner students had a higher rate of missing data in the 2020/21 school year than in the 2019/20 school year. In grades 4–6, students who were eligible for the national school lunch program had higher rates of missing

data in 2020/21 relative to 2019/20. In grades 4 and 6, students not in special education had a higher rate of missing data in the 2020/21 school year, but in grade 4, this was also true for students who were in special education. The study team included a sensitivity analysis that accounted for missing data in appendix C.

Table A5. Percentage missing fall Northwest Evaluation Association Measures of Academic Progress mathematics scores, by student group and grade in 2019/20 and 2020/21

Grade level and year	Percentage missing among student groups									
	National school lunch program		English learner students		Special education		Race/ethnicity			
	Not eligible	Eligible	No	Yes	Not in	In	White	Black	Hispanic	Another race or multiple races
Grade 3										
2019/20	14	19	17	16	15	25	13	32	15	17
2020/21	16	19	14	26[a]	17	26	10	24[a]	22[a]	17
Grade 4										
2019/20	12	16	14	15	12	24	11	23	12	15
2020/21	17	23[a]	21[a]	18	18[a]	31[a]	16	37[a]	16	17
Grade 5										
2019/20	11	14	12	15	11	24	10	20	11	14
2020/21	15	21[a]	18[a]	17	16	29	13	38[a]	13	16
Grade 6										
2019/20	11	15	13	16	11	24	10	22	11	14
2020/21	15	23[a]	19[a]	21	18[a]	29	14	38[a]	15	18
Grade 7										
2019/20	16	17	16	20	15	28	16	25	12	18
2020/21	16	22	19	20	18	29	14	41[a]	15	17
Grade 8										
2019/20	14	17	15	21	14	28	15	22	13	15
2020/21	15	21	18	20	16	31	12	40[a]	13	17

Note: Percentages of students with each characteristic are similar (differences of 1 percentage point or less) in both the mathematics and reading samples, so only the mathematics sample characteristics are presented here.

a. The difference between 2019/20 and 2020/21 is greater than 5 percentage points.

Source: Authors' analysis of data provided by participating districts and the Illinois State Board of Education.

Analysis

This section describes the methods used to answer each of the study's research questions.

Research question 1. To estimate changes in learning in mathematics and reading following the onset of the COVID-19 pandemic, the study team conducted an interrupted time series (ITS) analysis. ITS analysis entails estimating a regression model that includes time-based covariates. This approach takes advantage of data spanning multiple time points and accounts for a preexisting time trend in the outcome of interest leading up to an intervention or natural experiment—such as the COVID-19 pandemic—that affects the entire population at the same time (Kontopantelis et al., 2015).

The ITS analyses conducted for this study included four time points prior to the pandemic (the fall terms of 2016, 2017, 2018, and 2019) and one time point following the onset of the pandemic (the fall term of 2020). Separate ITS analyses were conducted for each grade level; each analysis included data on consecutive cohorts of same-

grade students (for example, observations of grade 4 students in the fall terms of 2016, 2017, 2018, 2019, and 2020). Separate ITS analyses were conducted for mathematics and reading.

The study team used the following multilevel regression model (model 1)—with students nested in schools nested in districts nested in time—to measure the relationship between the onset of the COVID-19 pandemic and students' fall NWEA MAP scores, while controlling for student, school, and district covariates and a time trend.

Model 1

Level 1 model: Student level

$$Y_{isdt} = \beta_{0sdt} + \beta_{1sdt}Time_{isdt} + \beta_{2sdt}PostCOVID_{isdt} + Y_{isdt-1}\beta_{3sdt} + \boldsymbol{Student}_{isdt}\beta_{4sdt} + \varepsilon_{isdt}$$

Level 2 model: School level

$$\beta_{0sdt} = \gamma_{00\mathrm{dt}} + \gamma_{01\mathrm{dt}}\boldsymbol{School}_{sdt} + r_{0sdt}$$

$$\beta_{1sdt} = \gamma_{10\mathrm{dt}}$$

$$\beta_{2sdt} = \gamma_{20\mathrm{dt}}$$

$$\beta_{3sdt} = \gamma_{30\mathrm{dt}}$$

Level 3 model: District level

$$\gamma_{00\mathrm{dt}} = \pi_{000t} + \pi_{001t}\boldsymbol{District}_{dt} + \sigma_{00dt}$$

$$\gamma_{10\mathrm{dt}} = \pi_{100t}$$

$$\gamma_{20\mathrm{dt}} = \pi_{200t} + \sigma_{20dt}$$

$$\gamma_{30\mathrm{dt}} = \pi_{300t}$$

Level 4 model: Time (years) level

$$\pi_{000t} = \rho_{0000}$$

$$\pi_{100t} = \rho_{1000}$$

$$\pi_{200t} = \rho_{2000}$$

$$\pi_{300t} = \rho_{3000}$$

Y is the fall test score for student *i* in school *s* in district *d* in year *t*, and Y_{isdt-1} is student *i*'s prior-year score in the same subject at time *t*–1. *Time* is a counter for the year that is zero-centered around the 2020/21 academic year, meaning that it takes the value of –4 in the first year of the data (2015/16), increases by 1 each year, and takes the value of 0 in 2020/21. *PostCOVID* is an indicator that equals 1 in the time period following the onset of the COVID-19 pandemic (fall 2020). ***Student*** is a vector of student demographic characteristics, ***School*** is a vector of school characteristics, and ***District*** is a vector of district characteristics. β_2 is the coefficient of interest and

captures the level change in the outcome Y (fall NWEA MAP scores in mathematics or reading) associated with the onset of the pandemic.[5]

After the coefficients were estimated, the study team calculated the model-adjusted predicted test scores for students enrolled in years prior to the COVID-19 pandemic and for students enrolled in fall 2020, fixing all other covariates at their means (for categorial variables, probability means were used). The model-adjusted test scores were then used to calculate learning changes from the pre-COVID-19 years to the COVID-19 year.

Research question 2. To measure how the estimated changes in learning in mathematics and reading differ for different types of students, the study team included interaction terms between *COVID* and student characteristics (model 2). A reduced-form version of model 1 with the interaction terms follows:

Model 2

$$Y_{isdt} = \beta_0 + \beta_1 Time_{isdt} + \beta_2 COVID_{isdt} + Y_{isdt-1}\beta_3 + \boldsymbol{Student}_{isdt}\beta_4 + \boldsymbol{School}_{sdt}\beta_5 + \boldsymbol{District}_{dt}\beta_6$$
$$+ \beta_7 PostCOVID_{isdt} * StudentCharacteristic_{isdt}$$
$$+ \sigma_{20dt} PostCOVID_{isdt} + \sigma_{00dt} + r_{sdt} + \varepsilon_{isdt}$$

β_7 captures the extent to which changes in student learning differ by the student characteristic of interest.

To measure how the estimated changes in learning in mathematics and reading differ for smaller and larger districts, the study team included interaction terms between *COVID* and district size (model 3). For simplicity while looking at interaction terms, the research team converted district size (a continuous variable) to a dichotomized variable using the sample average size to assign districts into the below-average group and the at-or-above-average group. A reduced-form version of model 1 with the district size interaction terms follows:

Model 3

$$Y_{isdt} = \beta_0 + \beta_1 Time_{isdt} + \beta_2 COVID_{isdt} + Y_{isdt-1}\beta_3 + \boldsymbol{Student}_{isdt}\beta_4 + \boldsymbol{School}_{sdt}\beta_5 + \boldsymbol{District}_{dt}\beta_6$$
$$+ \beta_7 COVID_{isdt} * LargerDistrict_{dt} + \sigma_{20dt} COVID_{isdt} + \sigma_{00dt} + r_{sdt} + \varepsilon_{isdt}$$

β_7 captures the extent to which changes in student learning differ for larger districts.

Conversion to days of instruction. For ease of interpretation, changes in NWEA MAP scores were converted into instructional time. This conversion was done by comparing the estimated change for a specific grade level and subject to the average growth observed in this subject among students entering this grade level. Average growth was determined using NWEA 2020 MAP assessment norms (Northwest Evaluation Association, 2020). For example, the average grade 3 fall score is 186.6 and the average grade 4 fall reading score is 196.7, meaning that the average growth between the start of grade 3 and the start of grade 4 is 10.1 points. The conversion assumed that these 10.1 points reflect 176 days of learning (the length of a full school year in Illinois). An estimated assessment score difference of -2 points is 20 percent of what grade 4 students gained since the start of grade 3, on average, and 20 percent of 176 days is 35 days. The average growth in NWEA MAP scores used in this conversion for each grade level and subject is in table A6.

[5] The study accounts for clustering of students in a time period and uses the mixed-effect model described here to take clustering by time into account. However, a mixed-effect model assumes a large number of clusters for a precise estimate of cluster-level variance. By contrast, an individual-level analysis with a small number of clusters—such as in the current study, which has a small number of years in which students nest—can lead to an inflated type I error rate (Leyrat et al., 2018), causing models to find statistically significant relationships where no relationships actually exist. Therefore, the model estimates should be interpreted with caution. Readers are advised not to focus on a specific point estimate or a specific standard error but rather on the direction of, and overarching patterns in, estimated learning changes.

Table A6. Growth in Northwest Evaluation Association Measures of Academic Progress average fall scores by grade level and subject

	Growth in NWEA MAP average fall score	
Grade level	Mathematics	Reading
Grade 3	13.5	14.3
Grade 4	11.1	10.1
Grade 5	9.6	7.8
Grade 6	5.6	5.7
Grade 7	5.5	4.0
Grade 8	4.7	3.8

NWEA MAP is Northwest Evaluation Association Measures of Academic Progress.
Source: Authors' calculations using Northwest Evaluation Association 2020 Measures of Academic Progress assessment norms.

References

Kontopantelis, E., Doran, T., Springate, D. A., Buchan, I., & Reeves, D. (2015). Regression based quasi-experimental approach when randomisation is not an option: Interrupted time series analysis. *BMJ, 350,* h2750.

Leyrat, C., Morgan, K. E., Leurent, B., & Kahan, B. C. (2018). Cluster randomized trials with a small number of clusters: Which analyses should be used? *International Journal of Epidemiology*, *47*(1), 321–331.

Northwest Evaluation Association. (2020). *2020 NWEA MAP growth normative data overview.* https://teach.mapnwea.org/impl/MAPGrowthNormativeDataOverview.pdf

Appendix B. Supporting analysis

This appendix presents estimates from the regression models (tables B1–B12). The appendix also presents figures (figures B1–B10) that illustrate estimated changes in learning across student groups in reading in all grades and mathematics in grade 3.

Table B1. Regression coefficients and standard errors from grade 3 mathematics and reading analyses

	Mathematics		Reading	
Predictor/covariate	**Coefficient**	**Standard error**	**Coefficient**	**Standard error**
Student-level characteristics				
COVID-19	-0.827	0.750	3.162	2.285
Black	-2.224***	0.280	-2.864***	0.339
Hispanic	-0.687***	0.198	-0.598*	0.242
Another race or multiple races	1.588***	0.211	0.454	0.256
Prior-year score	0.695***	0.005	0.683***	0.005
Student age	-0.882**	0.325	-0.419	0.396
Male	1.433***	0.113	-0.177	0.137
Eligible for the national school lunch program	-0.904***	0.157	-1.439***	0.191
English learner student	-2.692***	0.160	-5.114***	0.197
In special education	-4.473***	0.181	-6.228***	0.218
Time-level characteristics				
Year	0.074	0.213	-0.491	0.654
District-level characteristics				
District size	< 0.001	< 0.001	< 0.001	< 0.001
Percentage White	-0.028	0.024	-0.007	0.037
Percentage eligible for the national school lunch program	-0.036	0.023	-0.026	0.038
School-level characteristics				
Percentage White	0.009	0.021	-0.021	0.022
Percentage eligible for the national school lunch program	-0.037	0.019	-0.026	0.021
Percentage English learner students	0.016	0.011	0.003	0.013
Constant	80.342***	3.269	78.344***	4.760
Random effects	**Standard deviation**	**Variance component**	**Standard deviation**	**Variance component**
Time intercept	< 0.001	0.005	< 0.001	0.001
District intercept	1.326	0.248	5.196***	0.507
School intercept	1.527***	0.113	1.471***	0.123
Residual	7.573***	0.040	9.177***	0.048
N	18,557		18,550	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B2. Regression coefficients and standard errors from grade 4 mathematics and reading analyses

	Mathematics		Reading	
Predictor/covariate	**Coefficient**	**Standard error**	**Coefficient**	**Standard error**
Student-level characteristics				
COVID-19	-2.791***	0.629	0.801	1.035
Black	-2.125***	0.187	-2.875***	0.220
Hispanic	-0.571***	0.152	-0.470**	0.180
Another race or multiple races	1.305***	0.162	0.103	0.192
Prior-year score	0.810***	0.004	0.723***	0.004
Student age	-1.295***	0.213	-1.093***	0.252
Male	1.109***	0.090	-0.245*	0.106
Eligible for the national school lunch program	-1.056***	0.118	-0.848***	0.140
English learner student	-2.012***	0.138	-4.283***	0.167
In special education	-3.966***	0.147	-5.581***	0.174
Time-level characteristics				
Year	0.282	0.180	-0.099	0.297
District-level characteristics				
District size	< 0.001	< 0.001	< 0.001	< 0.001
Percentage White	-0.048*	0.019	-0.019	0.022
Percentage eligible for the national school lunch program	-0.027	0.019	-0.003	0.023
School-level characteristics				
Percentage White	0.017	0.015	0.007	0.016
Percentage eligible for the national school lunch program	-0.029	0.015	-0.037*	0.016
Percentage English learner students	0.010	0.008	0.008	0.010
Constant	66.035***	2.507	77.335***	3.065
Random effects	**Standard deviation**	**Variance component**	**Standard deviation**	**Variance component**
Time intercept	< 0.001	0.000	< 0.001	0.003
District intercept	1.238	0.153	2.401***	0.287
School intercept	1.480***	0.084	1.396***	0.098
Residual	7.118***	0.031	8.424***	0.038
N	26,124		26,160	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B3. Regression coefficients and standard errors from grade 5 mathematics and reading analyses

	Mathematics		Reading	
Predictor/covariate	**Coefficient**	**Standard error**	**Coefficient**	**Standard error**
Student-level characteristics				
COVID-19	-2.704***	0.598	-0.074	0.537
Black	-1.132***	0.189	-1.722***	0.218
Hispanic	-0.588***	0.147	-0.503**	0.170
Another race or multiple races	1.758***	0.161	0.639***	0.185
Prior-year score	0.890***	0.004	0.732***	0.004
Student age	-0.459*	0.203	-0.336	0.235
Male	0.470***	0.089	-0.309**	0.102
Eligible for the national school lunch program	-0.702***	0.115	-0.717***	0.134
English learner student	-1.889***	0.143	-3.913***	0.172
In special education	-2.738***	0.149	-4.544***	0.173
Time-level characteristics				
Year	0.017	0.171	-0.129	0.153
District-level characteristics				
District size	> -0.001	> -0.001	> -0.001	> -0.001
Percentage White	-0.059***	0.017	-0.021	0.017
Percentage eligible for the national school lunch program	-0.039*	0.016	0.005	0.016
School-level characteristics				
Percentage White	0.022	0.013	0.023	0.014
Percentage eligible for the national school lunch program	-0.033**	0.012	-0.043***	0.013
Percentage English learner students	0.042***	0.007	0.032***	0.008
Constant	41.497***	2.592	66.973***	2.889
Random effects	**Standard deviation**	**Variance component**	**Standard deviation**	**Variance component**
Time intercept	< 0.001	0.000	< 0.001	0.000
District intercept	1.237	0.157	1.006	0.172
School intercept	1.158*	0.077	1.209*	0.093
Residual	7.058***	0.031	8.171***	0.036
N	26,397		26,350	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B4. Regression coefficients and standard errors from grade 6 mathematics and reading analyses

	Mathematics		Reading	
Predictor/covariate	**Coefficient**	**Standard error**	**Coefficient**	**Standard error**
Student-level characteristics				
COVID-19	-2.780***	0.455	-1.086	0.647
Black	-1.650***	0.183	-1.963***	0.215
Hispanic	-0.960***	0.137	-0.506**	0.162
Another race or multiple races	1.014***	0.155	0.801***	0.182
Prior-year score	0.818***	0.004	0.729***	0.004
Student age	-0.513**	0.185	-0.841***	0.219
Male	0.248**	0.085	-0.333***	0.099
Eligible for the national school lunch program	-0.753***	0.110	-0.973***	0.130
English learner student	-1.598***	0.141	-3.359***	0.173
In special education	-3.006***	0.144	-3.860***	0.172
Time-level characteristics				
Year	-0.142	0.131	0.144	0.187
District-level characteristics				
District size	< 0.001	< 0.001	< 0.001	< 0.001
Percentage White	-0.045*	0.021	-0.027	0.025
Percentage eligible for the national school lunch program	-0.029	0.017	-0.028	0.021
School-level characteristics				
Percentage White	0.021	0.021	0.020	0.025
Percentage eligible for the national school lunch program	-0.038*	0.016	-0.018	0.019
Percentage English learner students	0.030**	0.009	0.024	0.013
Constant	54.141***	2.504	75.459***	2.966
Random effects	**Standard deviation**	**Variance component**	**Standard deviation**	**Variance component**
Time intercept	< 0.001*	0.000	< 0.001	0.001
District intercept	0.689	0.213	1.215	0.224
School intercept	1.136	0.109	1.279*	0.123
Residual	6.734***	0.030	7.919***	0.035
N	26,129		26,039	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B5. Regression coefficients and standard errors from grade 7 mathematics and reading analyses

	Mathematics		Reading	
Predictor/covariate	**Coefficient**	**Standard error**	**Coefficient**	**Standard error**
Student-level characteristics				
COVID-19	-3.029***	0.649	0.0548	1.514
Black	-1.265***	0.195	-2.018***	0.222
Hispanic	-0.685***	0.144	-0.708***	0.164
Another race or multiple races	1.361***	0.165	0.736***	0.186
Prior-year score	0.928***	0.004	0.740***	0.004
Student age	-0.860***	0.190	-0.953***	0.217
Male	-0.283**	0.090	-0.649***	0.102
Eligible for the national school lunch program	-0.460***	0.118	-0.771***	0.134
English learner student	-1.844***	0.154	-3.087***	0.182
In special education	-2.649***	0.156	-3.321***	0.179
Time-level characteristics				
Year	0.233	0.189	0.118	0.439
District-level characteristics				
District size	> -0.001	> -0.001	< 0.001	< 0.001
Percentage White	-0.059*	0.025	-0.040	0.039
Percentage eligible for the national school lunch program	-0.015	0.022	0.019	0.035
School-level characteristics				
Percentage White	0.032	0.025	0.012	0.035
Percentage eligible for the national school lunch program	-0.038	0.021	-0.062*	0.030
Percentage English learner students	0.012	0.018	0.015	0.025
Constant	39.587***	2.833	76.594***	3.684
Random effects	**Standard deviation**	**Variance component**	**Standard deviation**	**Variance component**
Time intercept	< 0.001	0.000	< 0.001	0.001
District intercept	1.397*	0.195	3.730***	0.339
School intercept	0.877	0.134	1.145	0.152
Residual	7.047***	0.034	8.020***	0.036
N	25,484		25,469	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B6. Regression coefficients and standard errors from grade 8 mathematics and reading analyses

	Mathematics		Reading	
Predictor/covariate	**Coefficient**	**Standard error**	**Coefficient**	**Standard error**
Student-level characteristics				
COVID-19	-2.370***	0.672	-1.287	2.265
Black	-1.190***	0.191	-1.517***	0.217
Hispanic	-0.349*	0.141	-0.406*	0.160
Another race or multiple races	1.415***	0.163	0.745***	0.182
Prior-year score	0.933***	0.003	0.766***	0.004
Student age	-0.473**	0.179	-1.166***	0.204
Male	-0.218*	0.088	-0.860***	0.100
Eligible for the national school lunch program	-0.755***	0.115	-0.639***	0.132
English learner student	-1.016***	0.160	-3.009***	0.187
In special education	-2.218***	0.156	-3.579***	0.178
Time-level characteristics				
Year	0.156	0.196	0.573	0.655
District-level characteristics				
District size	< 0.001	< 0.001	> -0.001	> -0.001
Percentage White	-0.048	0.030	-0.037	0.044
Percentage eligible for the national school lunch program	-0.012	0.025	-0.018	0.042
School-level characteristics				
Percentage White	0.019	0.031	0.026	0.034
Percentage eligible for the national school lunch program	-0.052*	0.025	-0.022	0.029
Percentage English learner students	-0.008	0.021	-0.021	0.025
Constant	33.674***	2.868	76.811***	4.316
Random effects	**Standard deviation**	**Variance component**	**Standard deviation**	**Variance component**
Time intercept	< 0.001	0.000	0.316	0.551
District intercept	1.261	0.357	5.608***	0.463
School intercept	1.301	0.202	1.025	0.151
Residual	6.954***	0.031	7.957***	0.035
N	25,701		25,930	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B7. Regression interaction term coefficients and standard errors from grade 3 mathematics and reading analyses

	Mathematics		Reading	
Interaction variable	Coefficient	Standard error	Coefficient	Standard error
Student characteristic				
Black	1.426*	0.661	0.328	0.827
Hispanic	0.761	0.470	-0.518	0.593
Another race or multiple races	1.398**	0.526	0.392	0.639
Eligible for the national school lunch program	0.460	0.381	-0.018	0.469
English learner student	3.055***	0.361	2.363***	0.440
In special education	2.296***	0.447	1.729**	0.536
District size				
Larger district	-0.215	1.060	-1.029	3.325
N	18,557		18,550	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Note: The table presents the coefficients and associated standard errors on the interaction of an indicator variable set equal to 1 for students who have the relevant characteristic and set equal to 0 for those who do not. Each subject-specific column includes estimates from five separately estimated regression models. Each model includes one set of interaction terms at a time.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B8. Regression interaction term coefficients and standard errors from grade 4 mathematics and reading analyses

	Mathematics		Reading	
Interaction variable	Coefficient	Standard error	Coefficient	Standard error
Race/ethnicity				
Black	0.999*	0.457	-0.597	0.542
Hispanic	0.054	0.365	-0.318	0.437
Another race or multiple races	0.862*	0.403	-0.178	0.477
Eligible for the national school lunch program	-0.780**	0.295	-1.067**	0.349
English learner student	0.934**	0.294	1.916***	0.348
In special education	1.859***	0.353	2.151***	0.414
District size				
Larger district	-0.068	0.869	-1.050	1.476
N	26,124		26,160	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Note: The table presents the coefficients and associated standard errors on the interaction of an indicator variable set equal to 1 for students who have the relevant characteristic and set equal to 0 for those who do not. Each subject-specific column includes estimates from five separately estimated regression models. Each model includes one set of interaction terms at a time.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B9. Regression interaction term coefficients and standard errors from grade 5 mathematics and reading analyses

	Mathematics		Reading	
Interaction variable	Coefficient	Standard error	Coefficient	Standard error
Student characteristic				
Black	0.793	0.460	0.570	0.522
Hispanic	-1.212***	0.356	-0.966*	0.411
Another race or multiple races	0.589	0.399	0.462	0.461
Eligible for the national school lunch program	-0.572*	0.285	-0.609	0.327
English learner student	-0.354	0.307	-0.342	0.360
In special education	1.123**	0.347	1.975***	0.400
District size				
Larger district	-0.407	0.814	-2.241**	0.697
N	26,397		26,350	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Note: The table presents the coefficients and associated standard errors on the interaction of an indicator variable set equal to 1 for students who have the relevant characteristic and set equal to 0 for those who do not. Each subject-specific column includes estimates from five separately estimated regression models. Each model includes one set of interaction terms at a time.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B10. Regression interaction term coefficients and standard errors from grade 6 mathematics and reading analyses

	Mathematics		Reading	
Interaction variable	Coefficient	Standard error	Coefficient	Standard error
Student characteristic				
Black	0.974*	0.424	-0.043	0.512
Hispanic	-0.050	0.326	-0.574	0.391
Another race or multiple races	1.513***	0.380	0.443	0.450
Eligible for the national school lunch program	0.076	0.266	-0.443	0.317
English learner student	-0.958**	0.310	-1.301***	0.367
In special education	0.741*	0.325	0.660	0.384
District size				
Larger district	0.186	0.651	-1.257	0.896
N	26,129		26,039	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Note: The table presents the coefficients and associated standard errors on the interaction of an indicator variable set equal to 1 for students who have the relevant characteristic and set equal to 0 for those who do not. Each subject-specific column includes estimates from five separately estimated regression models. Each model includes one set of interaction terms at a time.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B11. Regression interaction term coefficients and standard errors from grade 7 mathematics and reading analyses

	Mathematics		Reading	
Interaction variable	Coefficient	Standard error	Coefficient	Standard error
Student characteristic				
Black	0.771	0.467	-0.591	0.545
Hispanic	-0.087	0.347	-0.785	0.401
Another race or multiple races	1.603***	0.398	0.420	0.452
Eligible for the national school lunch program	-0.447	0.286	-0.640	0.328
English learner student	-0.606	0.325	-1.235***	0.371
In special education	-1.482***	0.343	-1.211**	0.390
District size				
Larger district	1.174	0.904	-2.588	2.184
N	25,484		25,469	

* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Note: The table presents the coefficients and associated standard errors on the interaction of an indicator variable set equal to 1 for students who have the relevant characteristic and set equal to 0 for those who do not. Each subject-specific column includes estimates from five separately estimated regression models. Each model includes one set of interaction terms at a time.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table B12. Regression interaction term coefficients and standard errors from grade 8 mathematics and reading analyses

	Mathematics		Reading	
Interaction variable	Coefficient	Standard error	Coefficient	Standard error
Student characteristic				
Black	-0.223	0.452	-0.809	0.520
Hispanic	-0.031	0.334	-0.475	0.385
Another race or multiple races	1.522***	0.394	0.318	0.452
Eligible for the national school lunch program	-0.130	0.274	-0.871**	0.315
English learner student	-1.656***	0.330	-1.988***	0.377
In special education	-1.472***	0.347	-0.484	0.398
District size				
Larger district	0.578	0.957	-2.466	3.217
N	25,701		25,930	

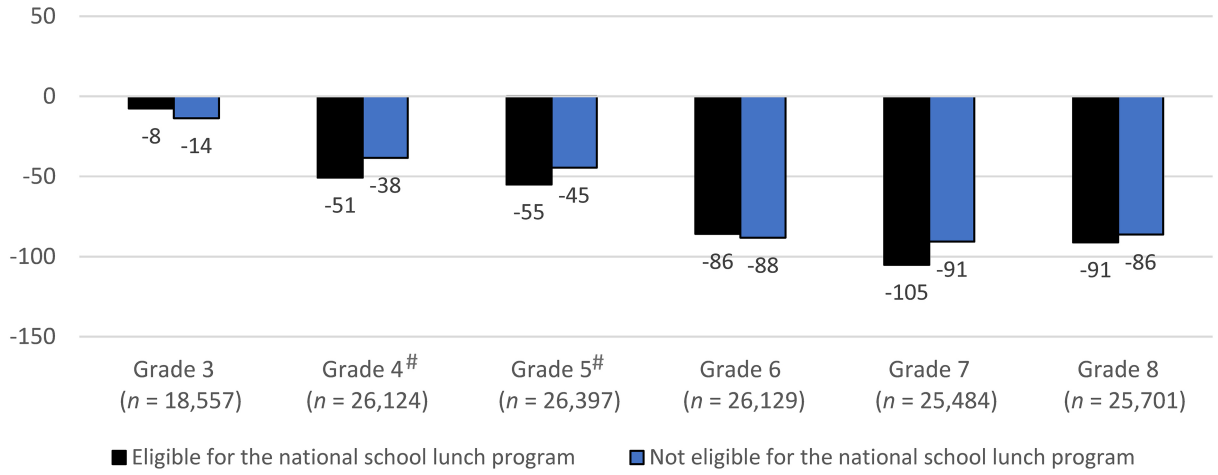
* Significant at $p < .05$; ** significant at $p < .01$; *** significant at $p < .001$.

Note: The table presents the coefficients and associated standard errors on the interaction of an indicator variable set equal to 1 for students who have the relevant characteristic and set equal to 0 for those who do not. Each subject-specific column includes estimates from five separately estimated regression models. Each model includes one set of interaction terms at a time.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B1. In grades 4 and 5, eligibility for the national school lunch program was negatively associated with changes in learning in mathematics, after adjusting for other factors, fall 2020

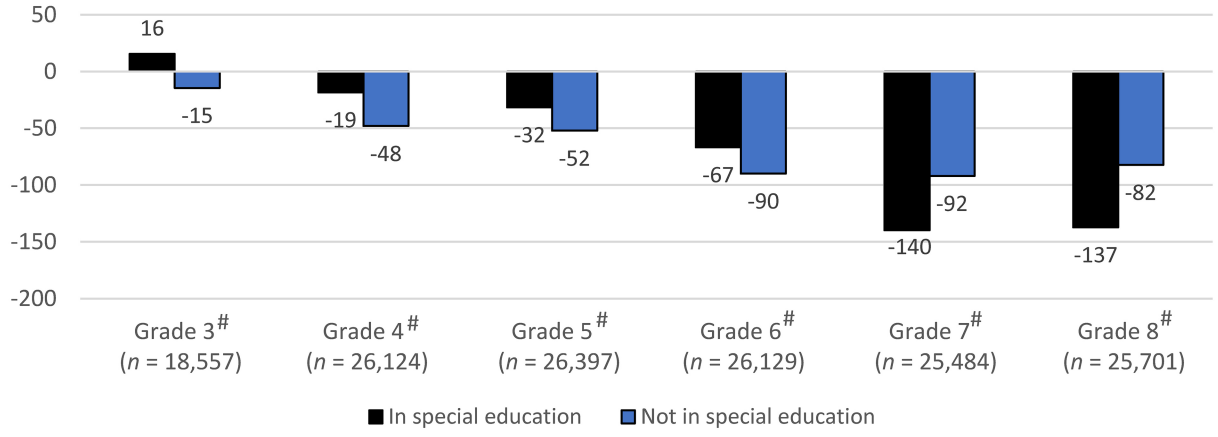
Estimated change in learning
(Days of instruction)



Differences between student groups were statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. All student groups experienced statistically significant changes in learning at $p < .05$ except grade 3 students who were, and were not, eligible for the national school lunch program.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B2. There was an association between being in special education and estimated change in learning in mathematics in all grade levels, after adjusting for other factors, fall 2020

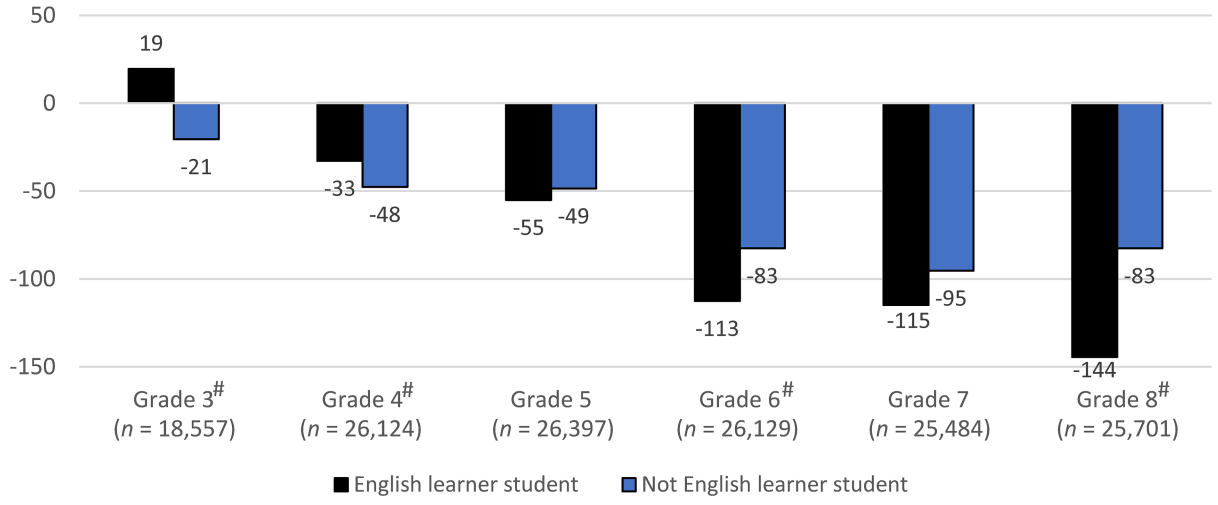
Estimated change in learning
(Days of instruction)



Differences between student groups were statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. All student groups experienced statistically significant changes in learning at $p < .05$ except grade 3 students who were, and were not, in special education.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B3. There were no consistent patterns in the associations between English learner status and changes in learning in mathematics, after adjusting for other factors, fall 2020

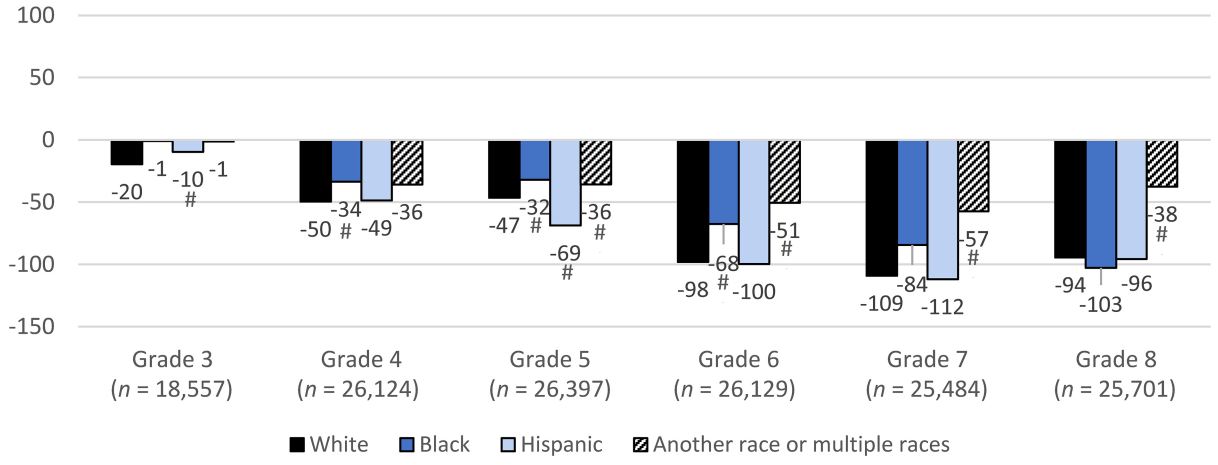
Estimated change in learning
(Days of instruction)



Differences between student groups were statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. All student groups experienced statistically significant changes in learning at $p < .05$ except grade 3 students who were, and were not, English learner students.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B4. In grades 3, 4, and 6–8, there was a positive association between being a student of another race or multiple races—relative to being a White student—and the estimated change in learning, after adjusting for other factors, fall 2020

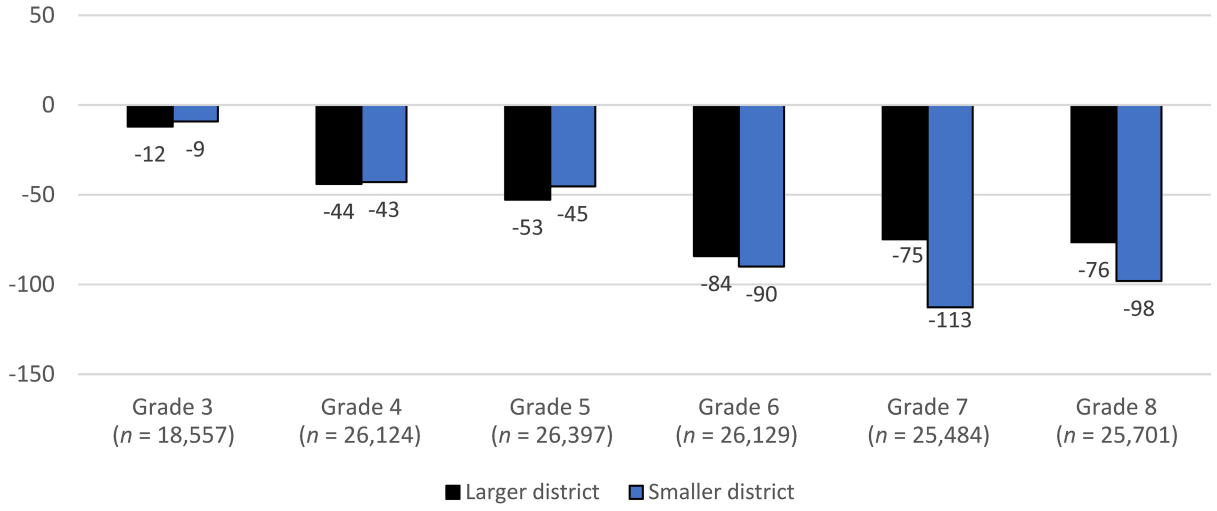
Estimated change in learning
(Days of instruction)



The difference between racial/ethnic group and White students was statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. Another race or multiple races includes students who are American Indian, Asian, Pacific Islander, or two or more races. All student groups experienced statistically significant changes in learning at $p < .05$ except students of another race or multiple races in grades 3–8, Black students in grades 3–5, White students in grade 3, and Hispanic students in grade 3.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B5. There was no association between district size and changes in learning in mathematics in any grade, after adjusting for other factors, fall 2020

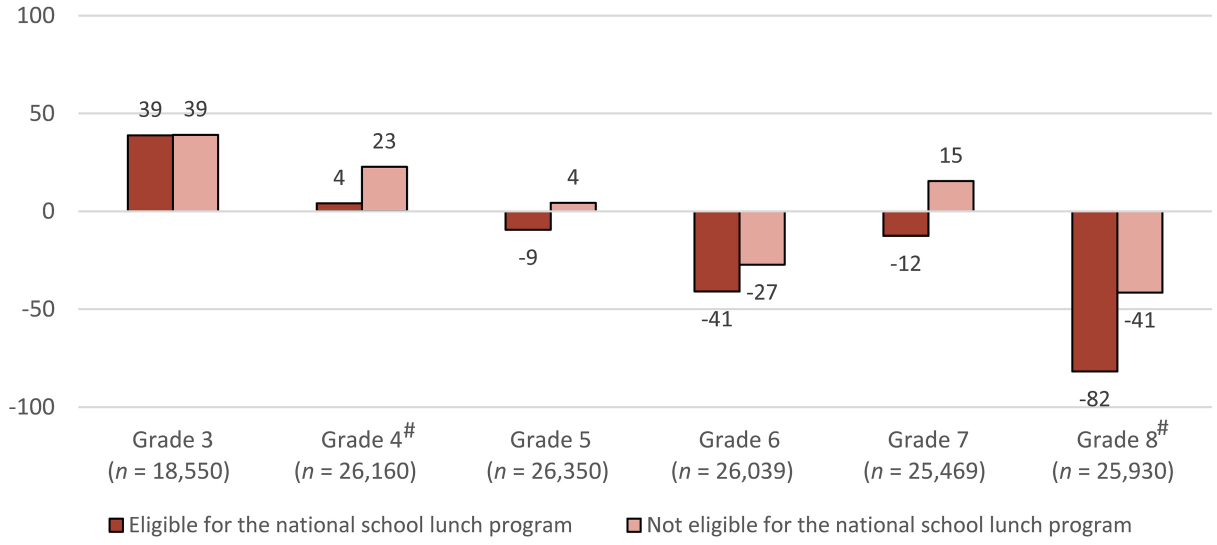
Estimated change in learning
(Days of instruction)



Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. Larger districts are those with total enrollment greater than or equal to the average district-level total enrollment; smaller districts are smaller than the average. The average district total enrollment is 3,449 for the grade 3 analysis and 3,998 for the grades 4–8 analyses. All student groups experienced statistically significant changes in learning at $p < .05$ except grade 3 students in larger and smaller districts. None of the differences between student groups is statistically significant.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B6. In grades 4 and 8, eligibility for the national school lunch program was negatively associated with changes in learning in reading, fall 2020

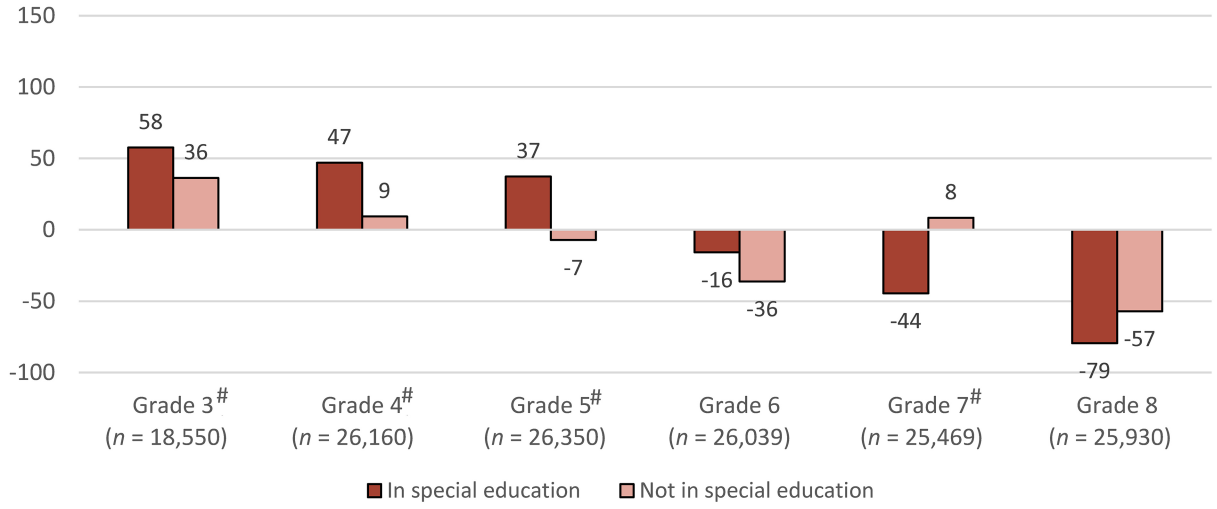
Estimated change in learning
(Days of instruction)



The difference between student groups was statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. None of the estimated changes in learning for individual student groups is statistically significant at $p < .05$.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B7. In grades 3–5 and 7, there was an association between being in special education and estimated changes in learning, after adjusting for other factors, fall 2020

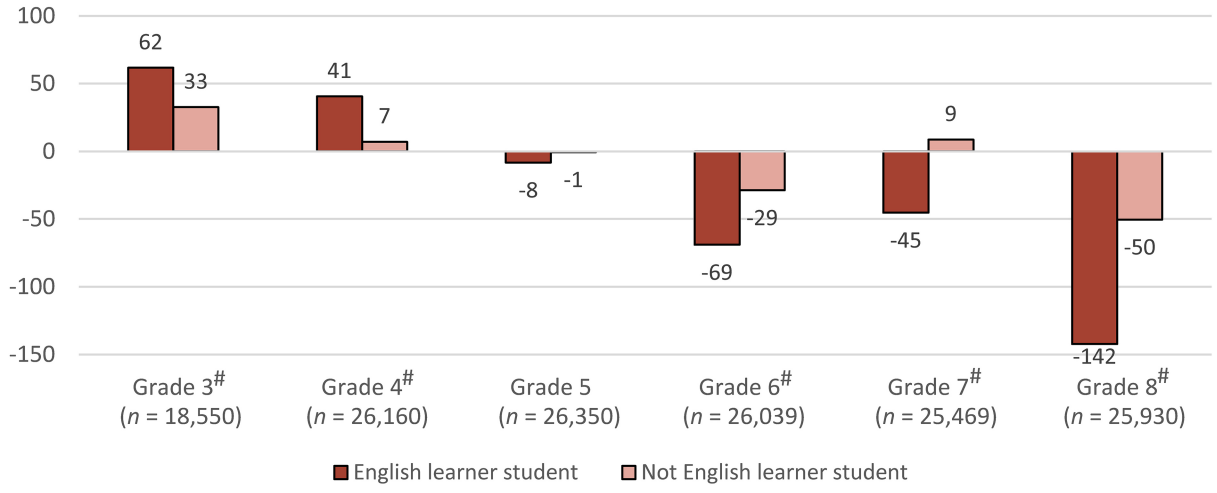
Estimated change in learning
(Days of instruction)



The difference between student groups was statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. None of the estimated changes in learning for individual student groups is statistically significant at $p < .05$.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B8. There was no consistent pattern in the association between changes in learning in reading and English learner status, after adjusting for other factors, fall 2020

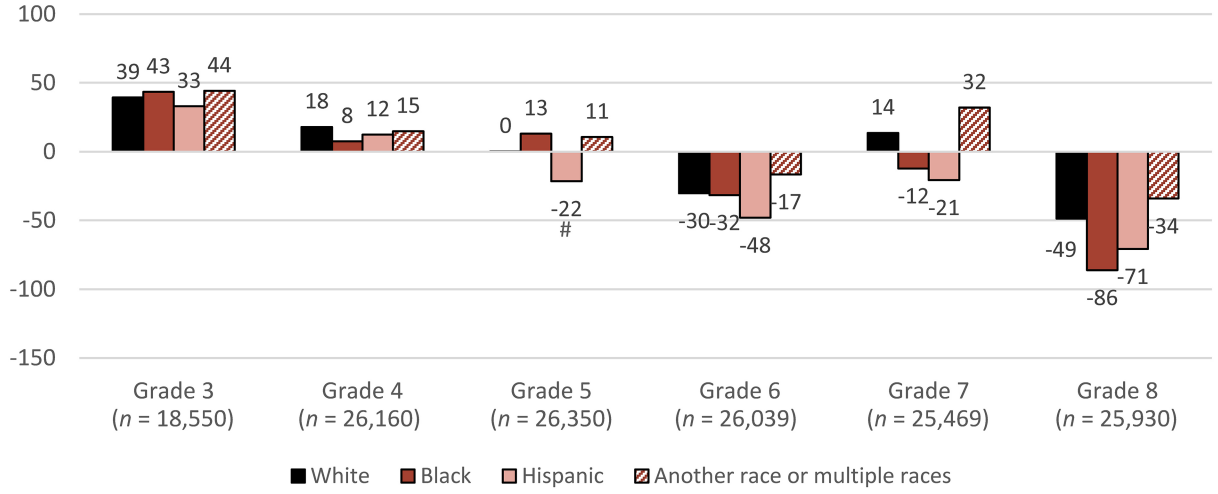
Estimated change in learning
(Days of instruction)



The difference between student groups was statistically significant at $p < .05$.
Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. None of the estimated changes in learning for individual student groups is statistically significant at $p < .05$ except for English learner students in grade 6.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B9. In grade 5, there was a negative association between being a Hispanic student—relative to being a White student—and changes in learning in reading, after adjusting for other factors, fall 2020

Estimated change in learning
(Days of instruction)



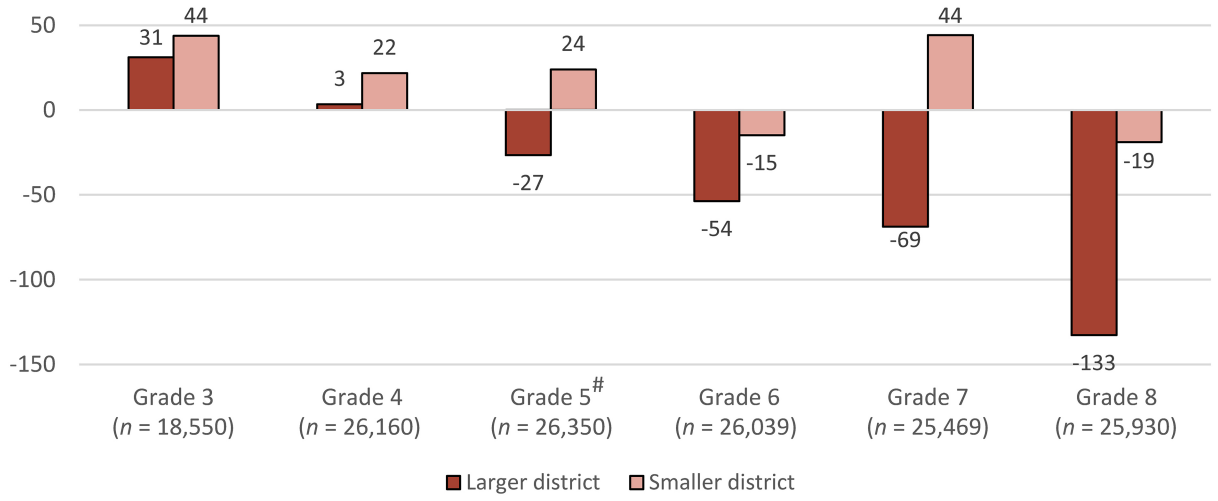
The difference between racial/ethnic group and White students was statistically significant at $p < .05$.

Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. Another race or multiple races includes students who are American Indian, Asian, Pacific Islander, or two or more races. None of the estimated changes in learning for individual student groups is statistically significant at $p < .05$.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure B10. In grade 5, there was a negative association between district size and estimated changes in learning in reading, after adjusting for other factors, fall 2020

Estimated change in learning
(Days of instruction)



The difference between student groups was statistically significant at $p < .05$.

Note: Estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. Larger districts are those with total enrollment greater than or equal to the average district-level total enrollment; smaller districts are smaller than the average. The average district total enrollment is 3,449 for the grade 3 analysis and 3,998 for the grades 4–8 analyses. None of the estimated changes in learning for individual student groups is statistically significant at $p < .05$.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Appendix C. Other analyses

The study team conducted supporting analyses that use different approaches to measure changes in student learning following the onset of the COVID-19 pandemic. The first supporting analysis was to calculate average (unadjusted) test scores for mathematics and reading for each grade level and in each time point. The purpose of this analysis was to describe the data without any statistical adjustments to account for student, school, and district characteristics or a time trend. The study team then recalculated these averages after adjusting for student, school, and district characteristics to illustrate the importance of the analytic strategy used to generate the main findings. The second supporting analysis replicated the analyses described in the main findings but used an alternative approach to conducting the conversion from estimated changes in Northwest Evaluation Association Measures of Academic Progress (NWEA MAP) scores to days of instruction that is described in appendix A. The third supporting analysis replicated the analysis described in the main findings but applied district- and student-level weights to improve balance between the sample and the statewide population. The fourth supporting analysis replicated the analysis described in the main findings but applied the student-level weights to test the sensitivity of the approach used for handling missing student-level data in the main findings.

Average mathematics and reading scores from fall 2016 to fall 2020

Although average unadjusted COVID-19–era fall mathematics achievement was lower than fall 2019 achievement in grades 4–8, the unadjusted COVID-19–era scores only appear to differ meaningfully from what would have been expected based on the time trend in grades 4 and 6 (figure C1). Yet, the main findings also showed significant unfinished learning in grades 5, 7, and 8, meaning that some of the results of this study are driven by inclusion of the covariates that adjust for changes in the composition of the students observed between the pre-COVID-19 and COVID-19–era time periods. Changes in the composition are illustrated in table C1. For example, the percentage of students who were English learner students in each grade level increased over time, and there is a meaningful increase (greater than 0.05 standard deviations) between 2019/20 and 2020/21 in average prior mathematics scores in grades 5 and 7.

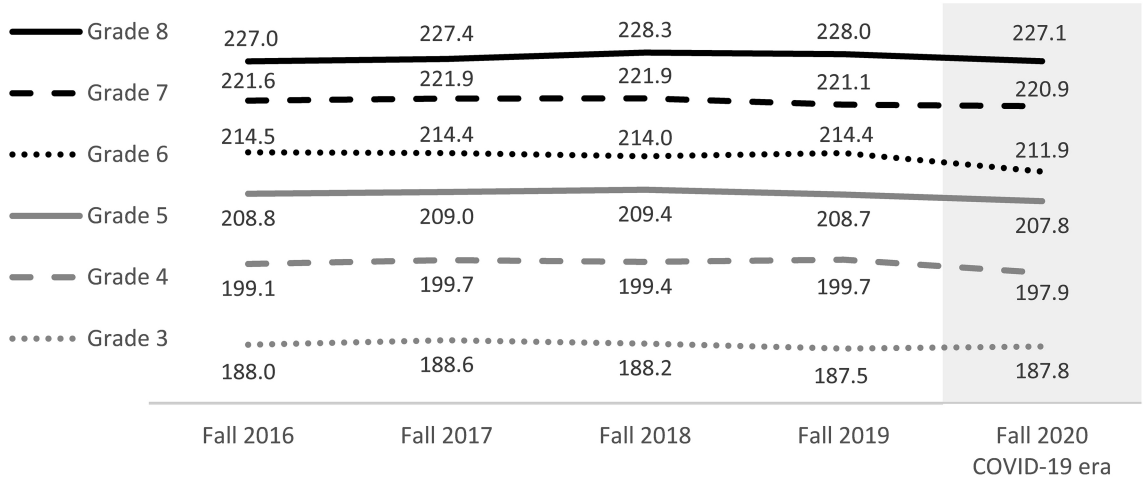
To further demonstrate this point, the study team recreated figure C1 but adjusted the averages for the student, school, and district characteristics that are included as covariates in model 1. After adjusting for student, school, and district characteristics, the trends in average scores show COVID-19–era achievement that is lower than what would have been expected based on the time trend in all grades, particularly so in grades 4–8, where the study findings are significant (figure C2).

For reading, the average unadjusted COVID-19–era fall achievement did not appear to differ from what would have been expected based on the time trend in any grade (figure C3). Furthermore, after adjusting for student, school, and district characteristics, the trends continue to show COVID-19–era fall achievement that is in line with what would have been expected (figure C4).

Furthermore, there are trends in average adjusted achievement over time. For example, average adjusted mathematics scores in grade 7 trend slightly upward between fall 2016 and fall 2019. One might expect this trend to continue, making the decline in average mathematics scores between fall 2019 and fall 2020, after adjusting for student, school, and district characteristics, even more noteworthy. The findings from these supporting analyses establish the need to adjust for student, school, and district characteristics as well as a time trend to provide better estimates of the relationship between the COVID-19 pandemic and student learning.

Figure C1. Average unadjusted mathematics scores in fall 2020 were lower than in fall 2019 in grades 4–8

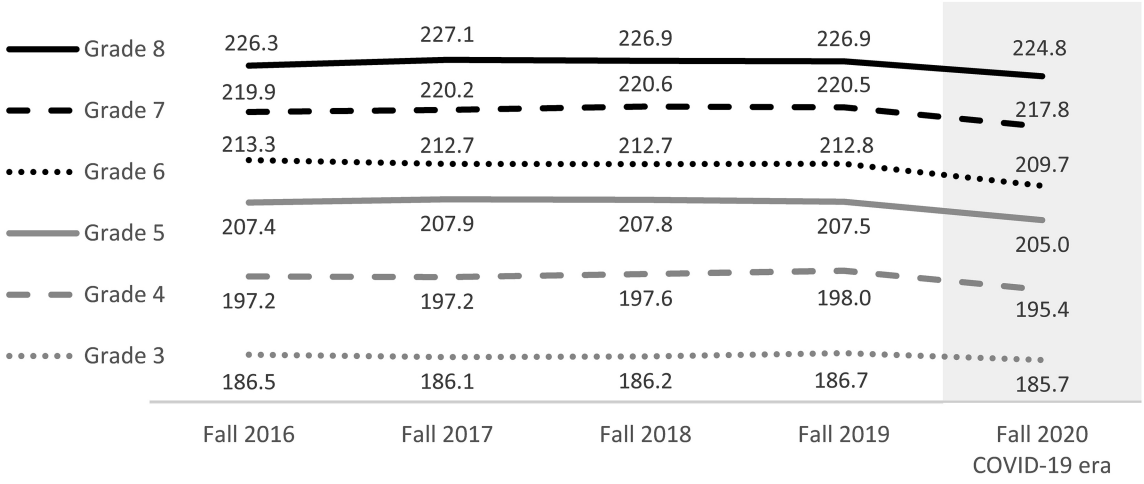
Northwest Evaluation Association Measures of Academic Progress score



Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure C2. Average adjusted mathematics scores in fall 2020 were lower than in fall 2019 in all grades

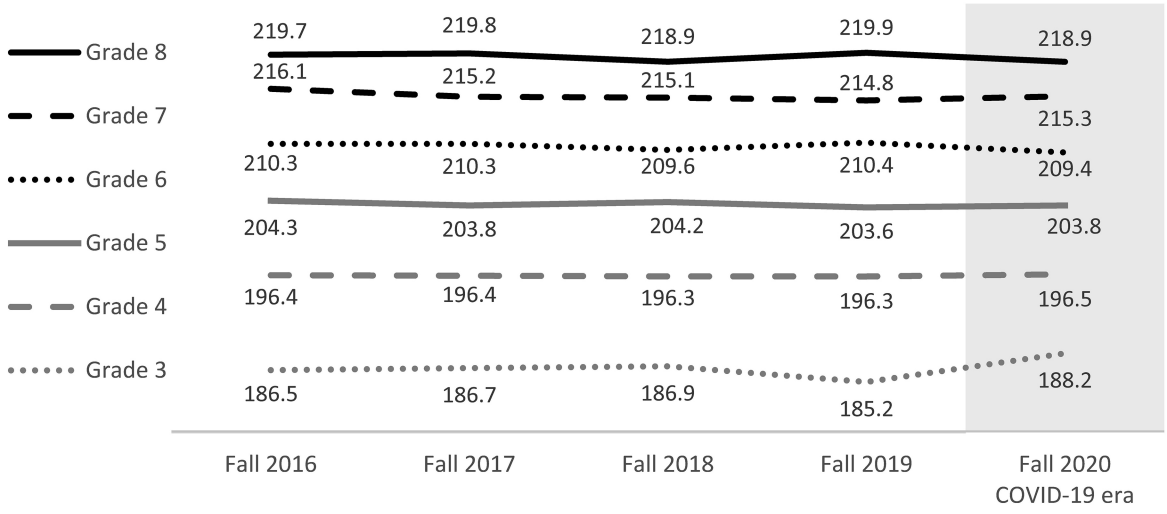
Adjusted Northwest Evaluation Association Measures of Academic Progress score



Note: This figure presents scores after adjusting for student, school, and district characteristics that are included as covariates in model 1.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure C3. Average unadjusted reading scores in fall 2020 were lower than in fall 2019 in grades 6 and 8

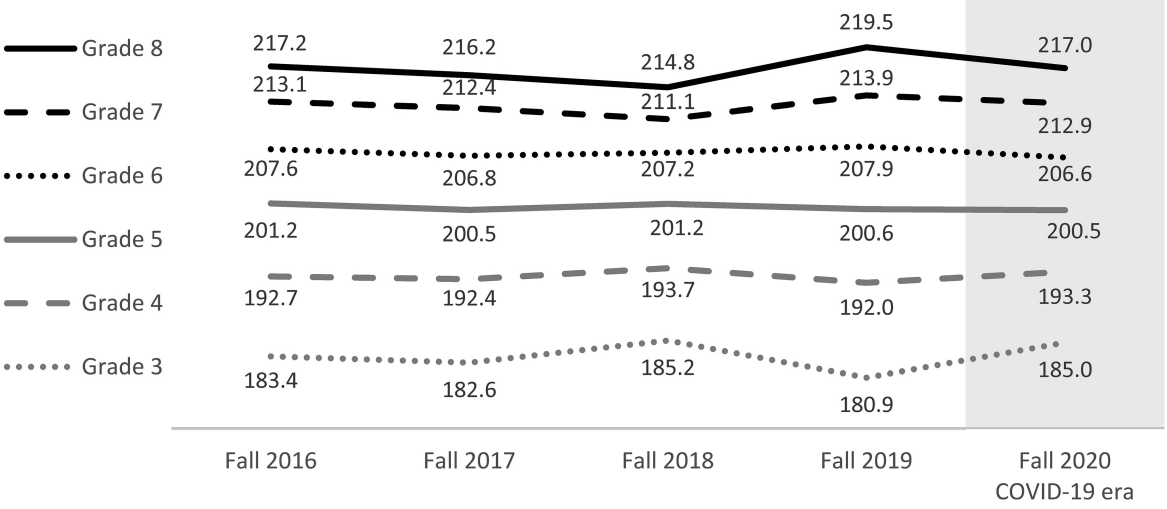
Northwest Evaluation Association Measures of Academic Progress score



Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure C4. Average adjusted reading scores in fall 2020 were lower than in fall 2019 in grades 5–8

Adjusted Northwest Evaluation Association Measures of Academic Progress score



Note: This figure presents scores after adjusting for student, school, and district characteristics that are included as covariates in model 1.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Table C1. Student characteristics by grade level and year

Grade level and year	Percentage of students who are:						Average prior mathematics score	Average prior reading score
	Eligible for the national school lunch program	English learner students	In special education	White	Black	Hispanic		
Grade 3								
2016/17	54	22	12	32	9	50	173.2	171.2
2017/18	52	21	12	32	10	48	174.2	171.4
2018/19	51	25	13	32	9	47	175.0	172.4
2019/20	50	32	13	33	11	45	174.7	173.2
2020/21	50	31	13	33	11	42	175.3	172.1
Grade 4								
2016/17	56	16	12	30	17	44	187.2	185.5
2017/18	55	18	13	31	16	42	188.0	186.5
2018/19	53	18	13	30	18	41	187.5	185.5
2019/20	54	22	13	31	17	40	188.0	186.6
2020/21	52	29	13	32	15	40	188.2	185.6
Grade 5								
2016/17	46	12	11	32	15	43	199.1	196.1
2017/18	55	16	12	30	16	43	199.2	196.7
2018/19	55	18	13	31	15	42	199.8	196.4
2019/20	53	18	13	31	17	41	199.5	196.4
2020/21	52	22	13	32	13	42	200.7	197.0
Grade 6								
2016/17	45	12	12	33	14	44	208.6	204.0
2017/18	45	12	11	33	15	42	208.9	204.4
2018/19	56	16	11	30	16	44	208.8	203.4
2019/20	54	18	13	31	15	42	209.5	204.2
2020/21	51	17	13	32	14	42	209.7	204.4
Grade 7								
2016/17	47	9	11	35	12	45	214.4	210.8
2017/18	44	11	11	33	14	42	214.7	210.8
2018/19	45	12	11	32	15	42	214.3	210.2
2019/20	56	16	12	30	16	43	213.9	209.4
2020/21	53	18	13	32	12	43	215.4	211.0
Grade 8								
2016/17	45	8	11	34	13	44	220.6	215.6
2017/18	45	9	11	34	14	43	221.2	216.3
2018/19	45	12	11	32	15	42	221.9	215.1
2019/20	45	12	11	32	16	41	221.8	214.9
2020/21	53	16	11	32	13	43	222.3	215.5

Note: Percentages of students with each characteristic are similar (differences of 2 percentage points or less) in both the mathematics and reading analytic samples, so the percentages from the mathematics analytic sample only are presented. Average prior mathematics and reading scores are calculated from the mathematics and reading analytic samples, respectively.

Source: Authors' analysis of data provided by participating districts and the Illinois State Board of Education.

Alternative method for converting estimated changes in Northwest Evaluation Association Measures of Academic Progress scores to days of instruction

The method used to convert estimated changes in NWEA MAP scores to days of instruction is described in appendix A. This conversion is done by comparing the estimated change for a specific grade level and subject to the average growth observed in this subject among students entering this grade level. The conversion uses growth in average scores between fall assessments in different grade levels, as determined using NWEA 2020 MAP assessment norms (Northwest Evaluation Association, 2020). The study team also calculated growth in average scores between fall assessments in different grade levels, as measured in the study sample, to compare the magnitudes of changes in learning between conversion methods.

The growth in average fall scores, for each grade level and subject, using these two methods are compared in table C2. The average growth measures within a subject and grade level are similar (differences are smaller than 0.05 standard deviations in scores for the grade level, as determined by the NWEA 2020 MAP assessment norms), except in grades 7 and 8 for mathematics and grade 7 for reading.

Table C2. Growth in average Northwest Evaluation Association 2020 Measures of Academic Progress fall scores, by sample, grade level, and subject

	Growth in average fall mathematics score		Growth in average fall reading score	
Grade level	NWEA 2020 MAP sample	Study sample	NWEA 2020 MAP sample	Study sample
Grade 3	13.5	13.1	14.3	13.5
Grade 4	11.1	11.7	10.1	10.4
Grade 5	9.6	9.4	7.8	7.5
Grade 6	5.6	5.4	5.7	6.2
Grade 7	5.5	7.1	4.0	5.0
Grade 8	4.7	6.2	3.8	4.5

NWEA MAP is Northwest Evaluation Association Measures of Academic Progress.
Note: The NWEA 2020 MAP sample column presents the growth in average scores from the NWEA 2020 MAP assessment norms. The study sample column presents the growth in average scores, as measured in this study's sample.
Source: Authors' calculations using NWEA 2020 MAP assessment norms and authors' analysis of data provided by school districts and the Illinois State Board of Education.

The estimated changes in learning converted to days of instruction when using these two different measures of average growth are presented in table C3. The columns labeled "NWEA 2020 MAP average growth" show the estimated changes in learning that were presented in the main report (figures 1 and 2). The estimated changes in learning are similar (differences between estimates reflect less than 0.05 standard deviations in test scores).

Table C3. Estimated changes in learning using alternative average growth measures, and differences in estimated changes, by grade level and subject

	Estimated change in learning in mathematics			Estimated change in learning in reading		
Grade level	NWEA 2020 MAP average growth	Study sample average growth	Difference (absolute value)	NWEA 2020 MAP average growth	Study sample average growth	Difference (absolute value)
Grade 3	-11	-11	< 1	39	41	2
Grade 4	-44	-42	2	14	14	< 1
Grade 5	-50	-50	1	-2	-2	< 1
Grade 6	-87	-91	4	-34	-31	3
Grade 7	-98	-75	22	2	5	3
Grade 8	-89	-67	21	-59	-50	10

Sensitivity analyses to address sample balance

The districts in the study sample differ from the population of districts serving students in grades 3–8 in Illinois. Furthermore, within the sample districts, some students are missing data. Together, this means that the districts and students in the study sample are not representative of all districts and students in Illinois. The study team tested the sensitivity of the main analyses to the inclusion of weights that were developed to improve the representativeness of the sample.

Weight development. To develop the weights, the study team used publicly available data from the National Center for Education Statistics (NCES) Common Core of Data (U.S. Department of Education, Institute of Education Sciences, NCES, n.d.) and the U.S. Department of Education's ED*Facts* (U.S. Department of Education, n.d.). From the Common Core of Data, the study team used the following data elements for each district: district size and the percentages of students who are in special education, English learner students, eligible for the national school lunch program, and from different racial/ethnic groups. From ED*Facts*, the study team used data on the percentages of students in each district who were proficient in English language arts and mathematics.

First, the study team created a set of district-level weights. These weights were created through a response propensity weighting approach. This approach estimated district response propensities (the propensity to have participated in the study) through a logistic regression model with response status as the dependent variable and district characteristics as the predictors. The predictors included district size; region; and the percentages of students who are in special education, English learner students, eligible for the national school lunch program, from different racial/ethnic groups, proficient in English language arts, and proficient in mathematics. All predictors were measured in the 2018/19 school year—the midpoint of the years included in the analysis and the most recent year from which all values are available. A separate set of weights was created for the grade 3 analyses, because the analytic sample for these analyses only includes 14 of the 17 participating districts.

The estimated district response propensities were divided into five weighting classes using quintiles of predicted response propensity. In each weighting class, the districts' weights were calculated as the number of all districts in the class divided by the number of responding districts in the class.

Second, the study team created a set of student-level weights to account for missing data within the responding districts. Again, the study team employed response propensity weighting. The approach was applied for each year and grade combination, to account for variation in the student population over time and between grade levels. Separate sets of weights were created for the mathematics and reading samples. In this model, individual students' responses or nonmissing status is the dependent variable, and student characteristics are the predictors. The predictors included a set of indicators for students who are in special education, students who are English learner students, students who are eligible for the national school lunch program, students who are American Indian, students who are Asian, students who are Black, students who are Hispanic or Latino, students who are Native Hawaiian/Pacific Islander, students who are two or more races, and students who are male. Similar to the district weighting, the estimated response propensities were divided into five weighting classes using quintiles of predicted response propensity, and the students' weights were calculated as the number of all students in the class divided by the number of students with nonmissing data in the class.

For the multilevel regression analyses, the district and student weights were applied at their corresponding levels. All schools were assigned a weight of one, assuming equiprobability sampling at the school level. For the descriptive statistics, the district weights and student weights were multiplied together.

The characteristics of the students in the unweighted and weighted (using both district- and student-level weights) analytic samples are shown in table C4 and compared to the characteristics of the state population in the same

grade. For simplicity, table C4 shows these characteristics from the 2018/19 school year. This period reflects the midpoint year in the study sample and the year that was used in calculating the district-level weights. For some characteristics and in some grade levels, the weighting approach improved balance between the sample and the statewide population. For example, the weighted sample percentage of students who are Hispanic is much more similar than the unweighted percentage to the population percentage. However, for some characteristics and in some grade levels, the weighting reduced balance. For example, the weighted sample percentage of students who are eligible for the national school lunch program is much less similar than the unweighted percentage to the population percentage.

Table C4. Student characteristics in the district- and student-level weighted analytic sample, 2018/19

	Percentage of students who are:						
Grade level and sample	Eligible for the national school lunch program	English learner students	In special education	White	Black	Hispanic	Another or multiple races
Grade 3							
Unweighted sample	51	25[a]	13	32[a]	9[a]	47[a]	11
Weighted sample	43[a]	13[a]	14	55[a]	12[a]	23	10
State population	52	19	15	47	18	26	10
Grade 4							
Unweighted sample	53	18	13	30[a]	18	41[a]	11
Weighted sample	42[a]	12[a]	13	55[a]	11	24	9
State population	52	19	15	47	17	27	10
Grade 5							
Unweighted sample	55	18	13	31[a]	15	42[a]	12
Weighted sample	41[a]	10	13	56[a]	11[a]	24	9
State population	52	15	15	47	17	27	9
Grade 6							
Unweighted sample	56	16[a]	11	30[a]	16	44[a]	11
Weighted sample	39[a]	8	13	57[a]	11	24	9
State population	52	10	15	47	17	28	9
Grade 7							
Unweighted sample	45	12	11	32[a]	15	42[a]	11
Weighted sample	39[a]	7	13	57[a]	11[a]	24	9
State population	50	8	15	47	17	27	9
Grade 8							
Unweighted sample	45	12	11	32[a]	15	42[a]	11
Weighted sample	37[a]	6	12	58[a]	11	23	8
State population	49	7	14	48	16	27	9

Note: The percentages of students with each characteristic are similar (differences of one percentage point or less) in both the mathematics and reading analytic samples, so the percentages from the mathematics analytic sample are presented.

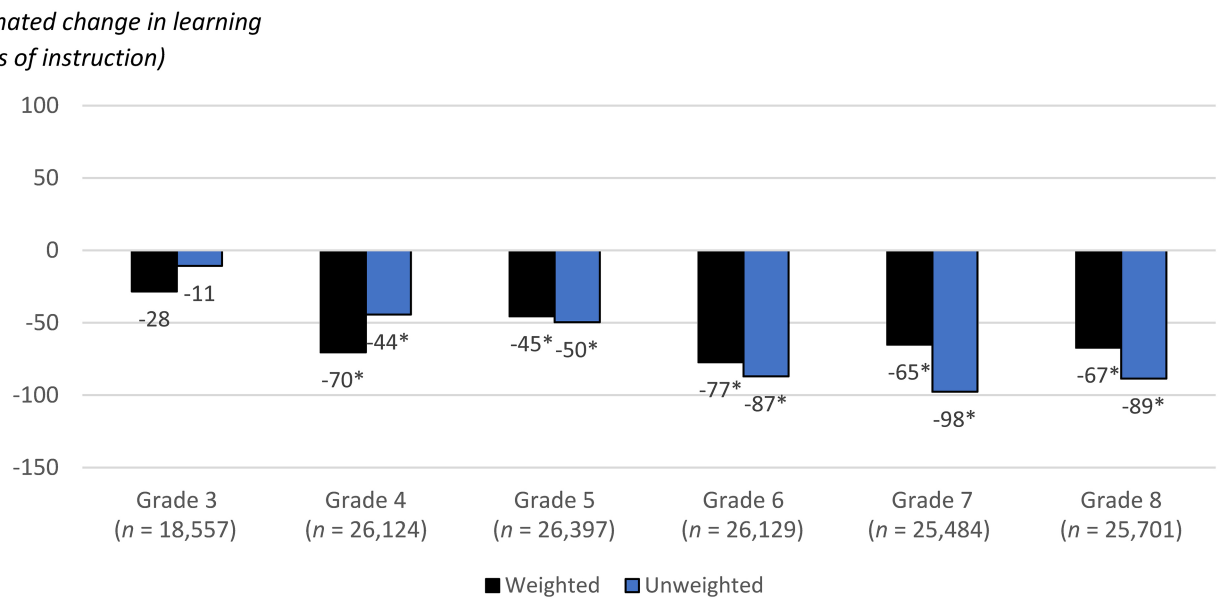
a. The difference between sample and statewide percentage > 5 percentage points.

Source: Authors' analysis of data provided by participating districts and the Illinois State Board of Education.

Weighted results. The results of the weighted and unweighted analyses are largely similar. Like the unweighted analyses presented in figure 1 and figure 2 in the main body of the report, students in grades 4–8 in the weighted

sample scored lower than expected in mathematics (figure C5). Students in grades 3–8 in the weighted sample did not score differently than expected in reading (figure C6). The magnitudes and directions in mathematics learning changes were similar in both the unweighted and weighted sample, except in grade 7 (where the difference reflected slightly more than 0.05 standard deviations in test scores). The direction and magnitude of the estimated change in learning for reading were similar between the unweighted and weighted samples for students in grades 4 and 8. In grades 3, 5, and 6, there were modest differences in the estimated changes in learning (reflecting slightly more than 0.05 standard deviations in test scores). In grade 7, there was a larger discrepancy between the estimated change in learning in the weighted sample (120 days) and the unweighted sample (2 days). The sample included a small number of districts (2 percent of the state population), and the weighting attempted to make these districts more representative of districts statewide. In doing so, the weights may have greatly increased the influence of an underrepresented district that had a large positive learning change in grade 7. Furthermore, applying weights leads to inflated standard errors, making it hard to detect changes in learning that are significantly different from zero. Together, these factors may explain the relatively large but statistically insignificant estimate in reading for grade 7 students in the weighted sample.

Figure C5. In the weighted sample, students in grades 4–8 scored lower than expected in mathematics, after adjusting for other factors, fall 2020



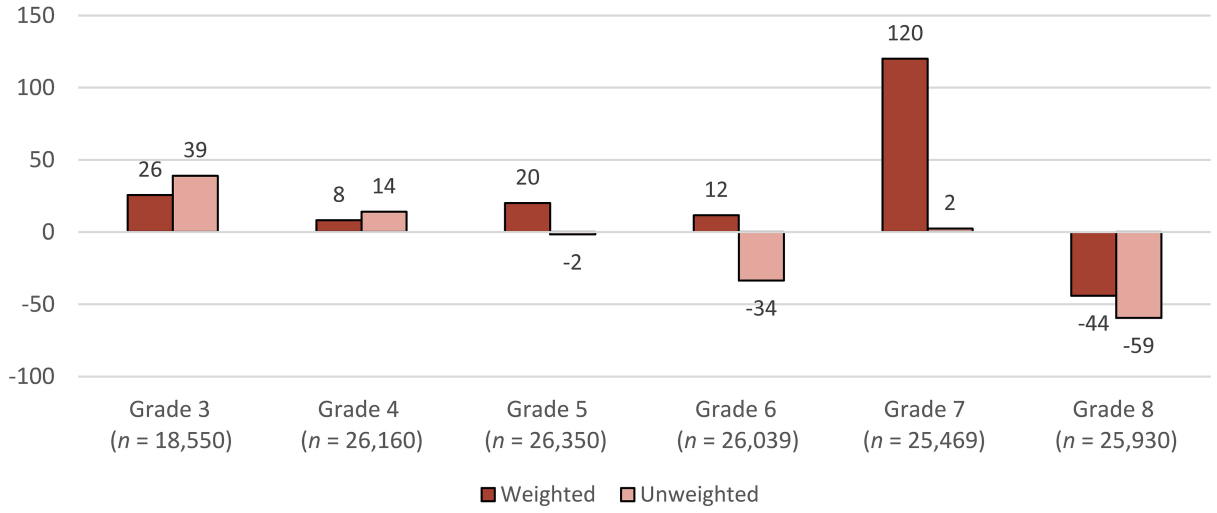
* The estimated change in learning was significant at $p < .05$.

Note: Weighted estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend and use district- and student-level weights. Unweighted estimates are the same as those presented in figure 1 in the main report. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days.

Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure C6. In the weighted sample, estimated changes in learning in reading were not significantly different from expected in grades 3–8, after adjusting for other factors, fall 2020

Estimated change in learning
(Days of instruction)



Note: Weighted estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend and use district- and student-level weights. Unweighted estimates are the same as those presented in figure 2 in the main report. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days. None of the estimated changes in learning is significant at $p < .05$.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Sensitivity analyses to address missing student-level data in sample districts

Within the sample districts, some students were missing data and were excluded from the analysis. For both mathematics and reading, each grade level excludes between 14 and 17 percent of enrolled students due to missing data. In the main analyses, the study team excluded students who were missing data and conducted a complete case analysis. The study team conducted a sensitivity analysis for dealing with missing student-level data within the participating districts by using a response propensity weighting approach to produce student-level weights that were representative of the district population. The characteristics of the students in the unweighted sample, the student-level weighted sample, and the participating district enrolled population are in table C5. Similar to table C4, table C5 shows these characteristics from the 2018/19 school year for simplicity. Comparisons of the characteristics of the analytic sample, the student-level weighted sample, and the enrolled population showed no differences greater than five percentage points.[6]

Table C5. Student characteristics in the student-level weighted analytic sample, 2018/19

	Percentage of students who are:						
Grade level and sample	Eligible for the national school lunch program	English learner students	In special education	White	Black	Hispanic	Another or multiple races
Grade 3							
Unweighted sample	51	25	13	32	9	47	11
Student weighted sample	53	26	14	31	11	47	11

[6] The characteristics compared include eligibility for the national school lunch program, English learner student status, special education status, and race/ethnicity.

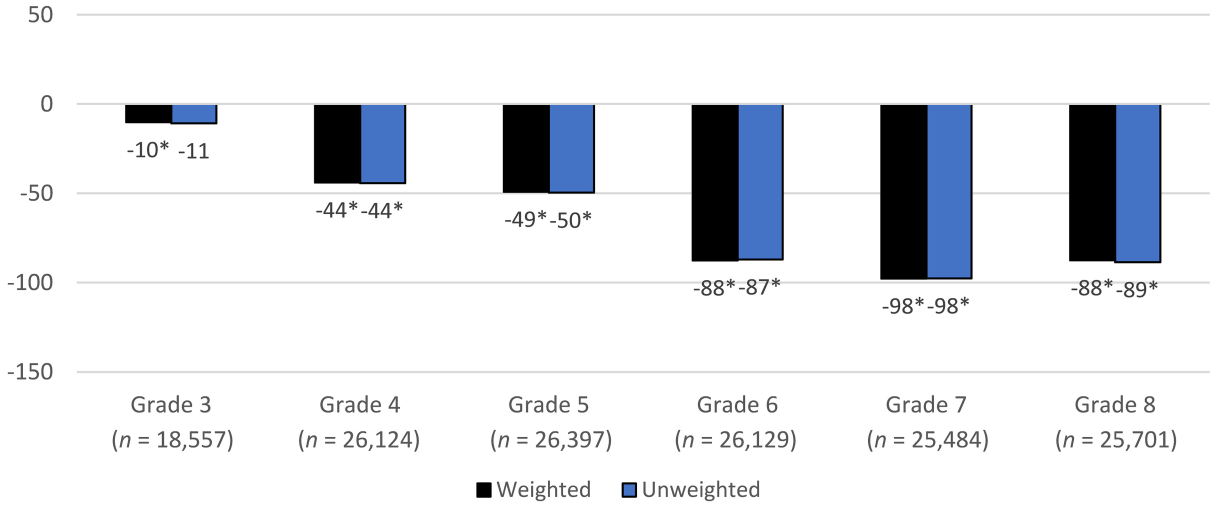
Grade level and sample	Percentage of students who are:						
	Eligible for the national school lunch program	English learner students	In special education	White	Black	Hispanic	Another or multiple races
Participating district population	52	25	14	31	12	47	11
Grade 4							
Unweighted sample	53	18	13	30	18	41	11
Student weighted sample	54	18	14	29	19	39	12
Participating district population	54	18	15	29	19	40	12
Grade 5							
Unweighted sample	55	18	13	31	15	42	12
Student weighted sample	56	19	14	29	17	41	12
Participating district population	56	19	15	30	17	41	12
Grade 6							
Unweighted sample	56	16	11	30	16	44	11
Student weighted sample	56	16	13	30	17	41	11
Participating district population	56	17	13	30	17	41	11
Grade 7							
Unweighted sample	45	12	11	32	15	42	11
Student weighted sample	46	13	12	32	17	40	11
Participating district population	46	12	13	32	17	40	11
Grade 8							
Unweighted sample	45	12	11	32	15	42	11
Student weighted sample	47	12	12	31	17	41	10
Participating district population	46	12	13	32	17	41	11

Note: The percentages of students with each characteristic are similar (differences of one percentage point or less) in both the mathematics and reading analytic samples, so the percentages from the mathematics analytic sample are presented. There were 14 participating districts for grade 3 and 17 participating districts for grades 4–8.
Source: Authors' analysis of data provided by participating districts and the Illinois State Board of Education.

Nonetheless, the study team tested the sensitivity of the approach to dealing with missing student-level data within the participating districts by using a response propensity weighting approach, in which student-level weights (described above) are applied in a replication of the main analyses. The aggregate mathematics and reading results are similar across these two approaches, with two exceptions. Applying the student-level weights results in a statistically significant reduction in learning in mathematics in grade 3 (figure C7) and a statistically significant reduction in learning in reading in grade 6 (figure C8). In both cases, the main findings were not statistically significant.

Figure C7. Students in grades 3–8 scored lower than expected in mathematics, after adjusting for other factors and applying student-level weights, fall 2020

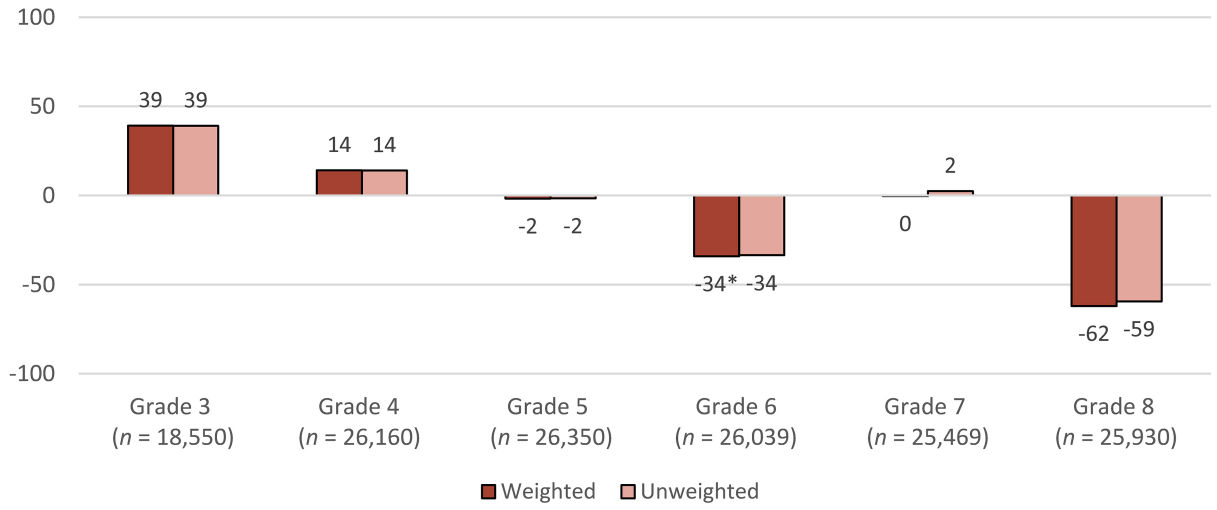
Estimated change in learning
(Days of instruction)



* The estimated change in learning was significant at $p < .05$.
Note: Weighted estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend and use student-level weights. Unweighted estimates are the same as those presented in figure 1 in the main report. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

Figure C8. Students in grades 3–5, 7, and 8 did not score differently from expected in reading, after adjusting for other factors and applying student-level weights, fall 2020

Estimated change in learning
(Days of instruction)



* The estimated change in learning was significant at $p < .05$.
Note: Weighted estimates are based on statistical models that adjust for student, school, and district characteristics and a time trend and use student-level weights. Unweighted estimates are the same as those presented in figure 2 in the main report. Estimated changes in learning are reported as the equivalent days of instruction. One year of instruction is equated to 176 days.
Source: Authors' analysis of data provided by school districts and the Illinois State Board of Education.

References

Northwest Evaluation Association. (2020). *2020 NWEA MAP growth normative data overview.* https://teach.mapnwea.org/impl/MAPGrowthNormativeDataOverview.pdf

U.S. Department of Education. (n.d.). *The ED*Facts *initiative.* https://www2.ed.gov/about/inits/ed/edfacts/index.html

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (n.d.). *Common Core of Data.* https://nces.ed.gov/ccd/