Declines in GPA During Remote Schooling Are Steepest for the Most Vulnerable Students

Benjamin Lira[a,1,2], **Maria E. Bartlett**[b,1], **Tim Kautz**[b], **and Angela L. Duckworth**[a]

This manuscript was compiled on August 5, 2024

Extensive research has documented steep declines in standardized achievement test scores during the COVID-19 pandemic. A smaller literature has attempted to isolate the influence of remote schooling in particular by associating regional differences in remote schooling practices with regional differences in test score declines. The current investigation addresses limitations inherent in the data available in prior research. We present quasi-experimental evidence that both substantiates the causal role of remote schooling on academic performance and, further, examines quarterly report card grades, a more direct measure of student engagement. Capitalizing on data from a large, demographically diverse school district—in which families were given the choice of remote versus in-person schooling during the 2020-21 academic year—we found that students who attended school remotely earned monotonically decreasing grades relative to their in-person counterparts, even after controlling for all time-invariant student characteristics. Declines were steeper for students with lower pre-pandemic report card grades, and for male, Black or Hispanic, and low-income students.

remote schooling | academic achievement | educational disparities

How does remote schooling influence academic engagement—and which students are influenced the most? It may seem like remote schooling is no longer a policy-relevant issue. However, in 2022-23, one in three large U.S. school districts reported offering students the choice of attending school remotely versus in person (1), and enrollment numbers have changed little since the pandemic (**?**). Moreover, the precedence and infrastructure of remote schooling options developed during the COVID-19 pandemic have created an alternative to in-person schooling on days when school might otherwise be canceled. For instance, on June 9, 2023, in response to poor air quality, the New York City Public Schools kept nearly 300,000 middle and high school students home to attend school remotely rather than in person (2). And districts across the United States offered remote schooling on election days in 2022 and 2023, when school facilities were used for polling (3). At present, days when students attend school remotely are assigned full credit toward state and federal requirements for school attendance. Additionally, there are strong incentives for providing virtual schooling options in rural districts, where geographical constraints and limited resources can make traditional schooling challenging (4).

There is strong evidence documenting unprecedented learning loss among elementary and middle school students during the COVID-19 pandemic (5). Across the United States, standardized achievement scores in math and reading declined sharply between 2019 and 2022 (6). Declines were especially steep for more vulnerable students (7). However, these analyses cannot disentangle the effects of remote schooling from other factors that may have contributed to learning loss (e.g., parent stress, teacher burnout, economic hardship, social unrest).

A smaller literature has examined the impact on standardized test scores of remote schooling in particular. An analysis of almost 5,000 districts in 12 states found that districts with fewer days of in-person instruction showed larger declines in third- through eighth-grade pass rates on state assessments (8). Districts with higher proportions of low-income, Black, or Hispanic students experienced disproportionately worse outcomes. Similarly, a national study found that third- through eighth-grade students who learned remotely for a greater proportion of the year showed greater learning loss relative to their pre-pandemic performance trends (9). Students with lower baseline test scores and Black and Hispanic students also saw greater declines relative to their pre-pandemic trends (5, 9). While suggestive, studies that have compared remote versus in-person schooling at the school or state level or where schooling modality changed across time for all students cannot rule out unmeasured third-variable regional confounds (e.g., regional differences in

Significance Statement

The decline of standardized test scores during the COVID-19 pandemic is well-established. A smaller literature has linked regional differences in these declines to regional differences in remote schooling—a practice still offered by many school districts. We present quasi-experimental evidence clarifying the relationship between remote schooling and learning loss. Additionally, we examine quarterly report card grades, which reflect not only academic skills but also student engagement. In a large, demographically diverse school district, students who attended school remotely earned monotonically decreasing grades relative to their in-person counterparts even after adjusting for all unchanging student characteristics. Declines were steepest for students with lower pre-pandemic report card grades and for male, Black or Hispanic, and low-income students.

Author affiliations: [a]University of Pennsylvania; [b]Mathematica Inc.

BL, MB, TK, and ALD conceptualized the study; BL, MB, TK, and ALD conducted statistical analyses; BL and ALD wrote the manuscript.

There are no competing interests to declare.

[1]B.L. contributed equally to this work with M.E.B.

[2]To whom correspondence should be addressed. E-mail: blira@sas.upenn.edu

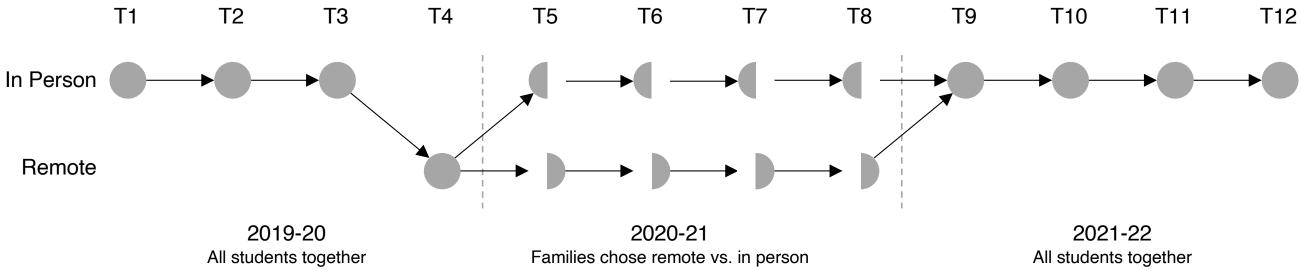

Fig. 1. During the first three-quarters of the 2019-20 academic year, all students in our sample attended in person. In the fourth quarter of 2019-20, all students attended school remotely. During the 2020-21 academic year, about half (44.8%) of students attended school in person, while the rest learned remotely. In the 2021-22 academic year, all students attended school in person.

the severity of the pandemic or internet connectivity) or temporal confounds (e.g., increases in the severity of the pandemic).

A second limitation of prior research concerns the outcome of standardized test scores. Standardized achievement tests administered annually by districts and states are designed to assess academic skills and knowledge, but they are a poor proxy for student behavior throughout the academic year (10). Report card grades, in contrast, reflect the aggregation of dozens of observations by students' teachers and have been shown to depend not only on students' academic mastery but also on demonstrated effort (11, 12). Not surprisingly, grades are often better predictors of consequential life outcomes, including high school and college graduation (13–15). One exception to this might be in the context of highly competitive admissions (i.e., Ivy-Plus colleges), where grades may be less predictive because so many applicants may have GPAs at the top of the scale (16). Thus, declines in report card grades commensurate with declines in standardized test scores would not only confirm learning loss but also hint at diminished student engagement as a potential mechanism.

In this investigation, we capitalized on longitudinal data from a large, diverse school district that offered high school students and their families the choice between fully remote and fully in-person schooling options for the 2020-21 academic year (see **Figure 1**). Specifically, we fit a panel regression with time and student fixed effects comparing, at each time point, the quarterly grade point average (GPA) of remote and in-person students. Including student fixed effects allowed us to control for any unmeasured time-invariant confounds at the student level. Including time fixed effects eliminated time trends common to all students. Our model included official student transcript data (N = 9,912) for three full academic years: In 2019-20, all students attended school in person for the first three quarters; during the fourth quarter, all students attended school remotely. In 2020-21, slightly more than half of families chose to send their children to school in person and the remainder kept their children home. In 2021-22, all students returned to in-person schooling.

Results

During the 2020-21 academic year, 44.8% (n = 4,439) of students opted to attend school remotely, and the remainder (55.2%; n = 5,473) opted to attend school in person. As shown in **Figure 2,** this choice covaried with student characteristics. In contrast to studies in which remote schooling policies were largely determined by policymakers, we did not find consistent evidence that students vulnerable to academic underachievement were more likely to opt into remote schooling. On one hand, students who attended school in person were more likely to be White or to speak English as a primary language at home. On the other hand, the choice to attend school in person was also more likely for students who were English language learners or in special education. Moreover, students who attended school remotely were more likely to be female or Asian. Most strikingly, the report card grades of students who opted for remote learning were higher, not lower, at each quarter of the prior (2019-20) academic year—and as shown in **Figure 3,** this gap showed a linear upward trend (b = 0.60, p < .001). Finally, students who were from low-income families (as indicated by eligibility for free or reduced-price meals), Black, or who spoke Spanish as a primary language at home were neither more nor less likely to attend school remotely versus in person. As shown in **Table S2,** this pattern of results was confirmed in probit models run as robustness checks.

Having established that students who would eventually attend school remotely were not underperforming at baseline, we proceeded with our main analyses. As shown in **Figure 3,** students who attended school remotely earned monotonically decreasing report grades, relative to their in-person peers in time 4, during each successive quarter of the 2020-21 academic year, even after controlling for all time-invariant student characteristics captured by the student-level fixed effects (implicitly including all those in **Figure 2**). By the fourth quarter, the grades of students learning remotely were d = 0.28 standard deviations (2.80 percentage points) below the grades of students learning in person. As a benchmark, this difference is more than nine times the median effect size of educational interventions funded by the U.S. Department of Education (d = 0.03 (17)). This gap disappeared when all students returned to in-person schooling in 2021-22. See **Table S3** for details.

In contrast to the larger impact of remote schooling on math (vs. reading) standardized test scores (9), the impact of remote schooling on GPA was similar for math and English language arts classes, see **Table S12**. This pattern of results was confirmed in robustness checks using available data on students in grades 7 and 8 during the 2020-21 academic year, see **Table S12**. We excluded these younger students from our main analyses because the school district was not able to provide complete information on their demographic characteristics, which were crucial for the moderation analyses we present next.

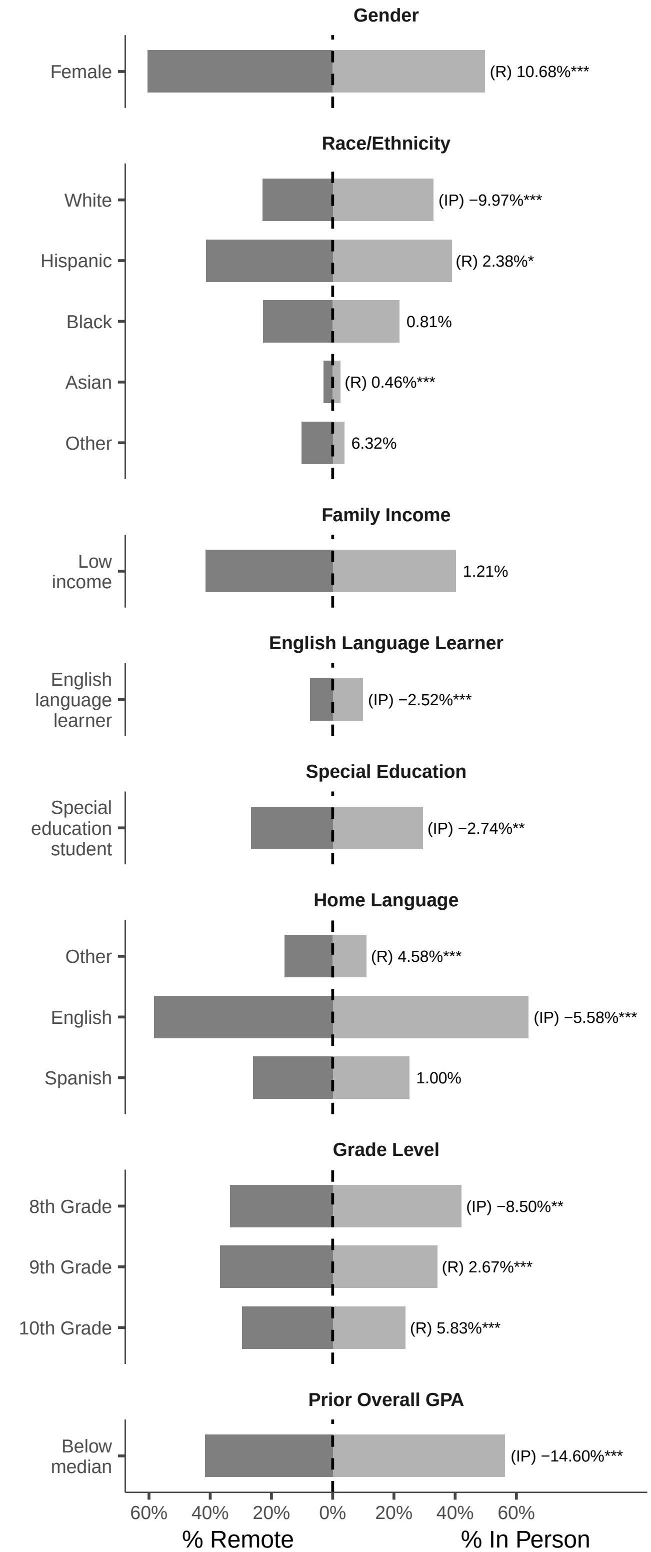

Fig. 2. Baseline differences between students who chose remote vs. in-person schooling. *p*-values come from Chi-square tests. Numbers to the right of each bar are the differences between proportions in remote and in-person schooling. Positive values indicate that this characteristic was more common in remote learning. Bars are labeled with R when the characteristic was significantly more common in remote schooling, and IP when the characteristic was significantly more common in in-person schooling. See Table S1 for details.
*** $p < .001$, ** $p < .01$, * $p < .05$

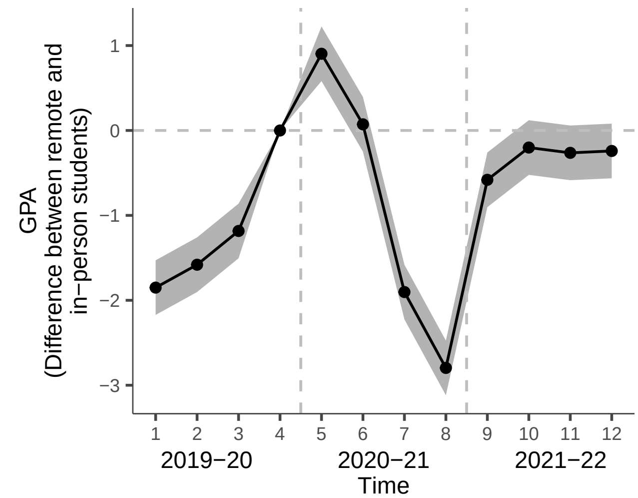

Fig. 3. Differences in Quarterly GPA (on a 100-Point Scale) Between Students Who Chose Remote (vs. In-Person) Schooling for the 2020-21 Academic Year. Shaded areas represent 95% confidence intervals. Estimates control for all time-invariant student characteristics through student fixed effects. Negative y-axis values indicate that students who learned remotely in the 2020-21 academic year earned lower report card grades than classmates who learned in person, relative to time 4. See Table S3 for details.

As shown in **Figure 4,** the declines in report card grades during remote (vs. in-person) schooling in 2020-21 were steeper for more vulnerable students. Specifically, declines (measured from Q1 to Q4 of 2020-21) were greater for students who were Black or Hispanic (compared to White or Other race, difference in declines = 2.54, $p < .001$; compared to Asian, difference in declines = 3.55, $p < .001$), male (difference in declines = 1.20, $p < .05$), from low-income families (difference in declines = 1.90, $p < .001$), or whose GPAs were below-median the prior year (difference in declines = 3.53, $p < .001$).

Likewise, these vulnerable groups disproportionately benefited from the district's decision to return all students to in-person schooling the following academic year. Specifically, recovery in GPAs (measured from Q4 of 2020-21 to Q1 of 2021-22) was greater for students who were Black or Hispanic (compared to White or Other race, difference in recoveries = 1.33, $p < .05$; compared to Asian, difference in recoveries = 3.48, $p < .001$), male (difference in recoveries = 1.63, $p < .001$), or whose GPAs were below the median in the 2019-20 school year (difference in recoveries = 2.66, $p < .001$).

As detailed in **Section C in Supplementary Materials,** GPA differences during 2020-21, and corresponding rebounds in 2021-22, were consistent across grade level, English language learner status, special education status, and home language.

Discussion

In a large and demographically diverse public school district, we examined quasi-experimental evidence for the impact of remote schooling on report card grades. Controlling for prior GPA, demographic covariates, and all other time-invariant student characteristics, high school students whose families opted for remote (vs. in-person) schooling during the 2020-21 academic year earned progressively lower report card grades each quarter of the 2020-21 academic year relative to students who opted for in-person schooling. The following year, when all students in the district returned to in-person schooling, this gap disappeared. More vulnerable students (i.e., Black, Hispanic, male, underachieving, and low-income students) were not systematically more likely to opt into remote schooling, but they suffered greater GPA declines when they did.

Given the continued post-pandemic decline in standardized achievement test scores nationwide, why did report card grades rebound immediately when students who had been learning remotely returned to in-person schooling in fall 2021 (18, 19)? And why did we find comparable results for report card grades in math and English language arts courses, whereas math test scores dropped more than reading scores during the pandemic (20)? These asymmetries underscore differences between test scores and report card grades. Standardized achievement tests are designed to provide an objective, apples-to-apples comparison of math and reading skills, whereas report card grades serve an additional purpose: an index of and an incentive for students to demonstrate engagement in classwork and homework, as judged by their own teachers. The discretion of teachers to award students higher or lower grades depending on perceived effort may also explain why test scores continue to drop post-pandemic while report card assigned by teachers nationwide are instead increasing (21).

Without randomly assigning students to remote versus in-person schooling, causal inferences about the impact of remote schooling are not warranted. However, concerns that selection bias accounts for the association between remote schooling and poor report card grades are somewhat mitigated by the observation that students who eventually chose remote schooling were actually outperforming their classmates prior to the 2020-21 academic year. That is, the parallel trends assumption did not hold in our data in a way that would have likely understated the negative effects of remote schooling. If the trend during the 2019-2020 school year held during the 2020-2021 school, remote schooling could have accounted for about a 5-point reduction in the GPA of students who attended remotely by the end of the 2020-21 school year, rather than a 2.80 point reduction.

To further explore the role of performance trends pre-pandemic, we used a synthetic differences-in-differences approach that accounted for student fixed effects but also weighted the remote students to have similar pre-trends as the in-person students (22). As shown in **Table S12,** comparing the outcomes for remote students to this weighted synthetic control of in-person students produces similar results, providing evidence that the apparent difference in pre-trends does not affect the conclusions.

Our main specification accounts for bias introduced by stable individual differences and general time trends, but is unable to account for time-varying-individual confounds. Any stable individual difference that might covary with the choice between remote or in-person schooling was captured by individual fixed effects. This includes all differences included in **Figure 2**. Any time trends, such as the pandemic getting progressively worse is also captured by period fixed effects. That only leaves the possibility of individual-time-varying confounds. For example, over time, schools and teachers

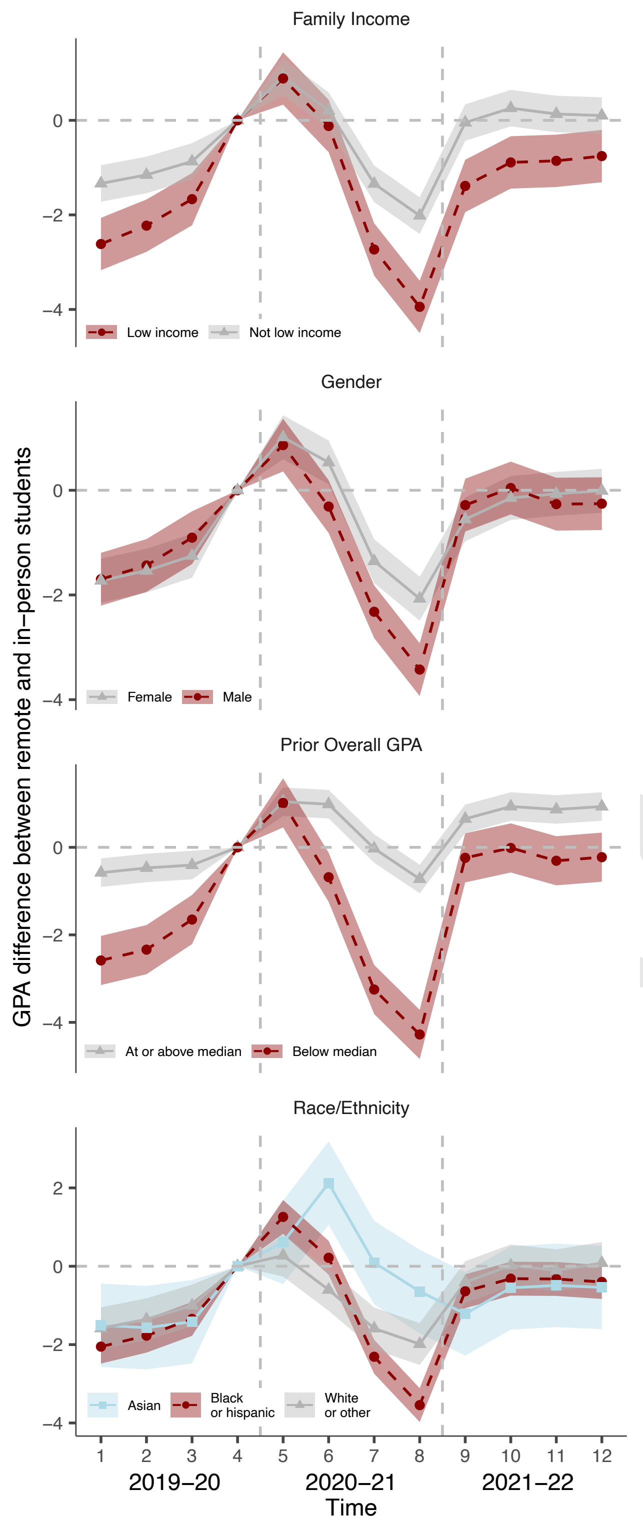

Fig. 4. Differences in GPA (on a 100-Point Scale) Between Students who Chose Remote vs. In-Person Schooling for the 2020-21 Academic Year Were Greatest for Low-Income, Lower-Achieving, Black or Hispanic, and Male Students. Shaded areas represent 95% confidence intervals. Students were considered to be from low-income families if they qualified for free or reduced-price meals. Negative y-axis values indicate that students who learned remotely in the 2020-21 academic year earned lower report card grades than classmates in the same school who learned in-person, relative to time 4.

may have adapted to remote schooling differently, family dynamics might have changed, or access to resources might have changed. These are unlikely to confound our results, given that they are likely uncorrelated with the decision to attend school remotely, over and above their levels at baseline. One notable exception is that, over time, students' mental health might have deteriorated more rapidly in remote as opposed to in-person schooling. Importantly, this would operate as a mediator rather than a confound. That is, part of the reason why remote schooling was detrimental to students' grades was because it influenced students' mental health. Prior evidence suggests this might very well be the case (23)

Another threat to causality in the present research is that given how schooling changed for all students during the COVID-19 pandemic, it is difficult to know whether our effects were driven by decreases for remote students, or improvements for in-person students. Thus, further research is needed to compare remote schooling to more naturalistic versions of in-person schooling. While not conclusive, our data suggests that after the initial decrease at the beginning of the 2020-21 school year, remote student achievement stagnated, while in-person students improved. See **Figure S1** in Supplement.

Three other limitations of this investigation are worth highlighting. First, we can only speculate as to why certain families choose one schooling modality over the other. We venture to guess that families were influenced by multiple factors, including prevailing gender norms (perhaps inclining families to keep girls at home), concerns about physical safety (inclining parents to keep their ninth graders at home more than their 10th or 11th graders), and expectations of academic independence (inclining students with higher GPAs to choose remote schooling).

Second, standardized test score data were not available for the high school students in our sample. This is not unusual, since district and state-mandated standardized testing is more common in elementary and middle school students in the United States. Nevertheless, this limitation makes it impossible to directly compare the changes in report card grades to changes in standardized test scores. Grades are also multiply determined, making it difficult to know how the observed effects could be partitioned into teacher behavior (e.g., grade inflation), student behavior (e.g, effort), and/or student skill development (24).

Finally, the generality of our findings cannot be assumed (25). The data in this study are from a single U.S. school district, and we cannot be certain that our findings would replicate in other school districts or countries, nor that the effects of remote schooling would be identical in a non-pandemic era. For instance, it may be that occasional remote schooling (e.g., for weather emergencies) is less deleterious than remote schooling over more extended periods. Likewise, we cannot know whether our observations of adolescents in high school—who have heretofore been understudied—would replicate in a sample of younger learners.

Notwithstanding these limitations, our results are consistent with a sober picture of remote schooling. Taken together with separate research showing that academic engagement, as well as emotional and social well-being, suffer during remote schooling (26), we recommend preserving the option of in-person schooling whenever possible, for the benefit of all students and especially the most vulnerable.

Materials and Methods

Participants and Procedure. All data for this investigation were collected from Orange County Public Schools in Florida—one of the 10 largest school districts in the United States—by Character Lab Research Network (CLRN), a national, nonprofit consortium of school partners committed to advancing scientific insights that help adolescents thrive. From school records, we obtained and averaged report card grades (on a 100-point scale) to create quarterly GPAs. To the best of our knowledge, grading policies for remote and in-person modalities were identical and did not change across the 3-year period of this investigation.

Our main sample comprised $N = 9{,}912$ students who were in grades nine, ten, and eleven in the 2020-21 academic year, and for whom transcript data were available during the prior (2019-20) and subsequent (2021-22) academic years. In spring 2020, families were offered the choice of attending school remotely or in person for the fall. Changing status within a quarter was not allowed. Remote (vs. in-person) status was a binary indicator obtained from a Character Lab Research Network student survey administered during the fourth quarter of 2020-21. In this district, course content for remote schooling was closely matched to in-person schooling. Both remote and in-person students could be assigned to the same teacher for synchronous class instruction. Remote students adhered to the same bell schedule as in-person students. A majority of teachers in the district (77%) taught both in-person and remote students.

As a robustness check, we repeated analyses with students who were in middle school in 2020-21, for whom certain 2019-20 demographic variables (e.g., status as English language learner), 2019-20 Quarter 1 grades, and school ID were missing in data received from the district. See **Table S12** for details.

Analytic Strategy. We used an event study methodology to estimate the difference in GPAs of students attending school remotely versus in person, relative to time t=4. Specifically, we estimated a fixed effects panel regression which controls for all time-invariant person-specific differences. Eliminating this variation was important because, as shown in **Table S1,** students whose families chose remote schooling were not equivalent at baseline. Specifically, we estimated the following model

$$y_{ist} = \alpha_i + \sum_{t=1}^{12} \psi_t + \left(\sum_{t=1}^{12} \lambda_t \times Remote_{is} \right) + \epsilon_{ist},$$

where y_{ist} is the GPA of student i in school s during time period t; α_i is a student fixed effect; ψ_t represents changes in each time point for the group of students who did not attend remote schooling; λ_t represents our parameters of interest, the difference in GPA between in-person students and students who were remote during the 2020-21 academic year; $Remote_{is}$ is an indicator for whether a student attended school in person during the 2020-21 academic year; δ_i ; and ϵ_{ist} is an error term. We use time t=4 as the reference period, so the estimated differences are relative to the difference between remote and in-person students during the period before remote schooling was optional. Standard errors were clustered at the individual level and were calculated using bootstrapping ($b = 500$).

To examine moderation effects, we fit the same fixed effects model within each demographic subgroup. For baseline GPA, we dichotomized the variable (students at or above the median vs. students below the median) and calculated the estimated differences within each subgroup. Additionally, we estimated the difference in covariate-adjusted declines in GPA for remote versus in-person students from Q1 to Q4 of the 2020-21 academic year and the corresponding differences in covariate-adjusted recoveries from Q4 of the 2020-21 academic year to Q1 of the 2021-22 academic year.

We tested whether these declines and recoveries differed between subgroups. We used bootstrapping to estimate the variation in the estimation of the demographic differences in declines and recoveries.

ACKNOWLEDGMENTS. We thank the CharacterLab Research Network and Orange County Public School System.

1. C Pangenigan, B Dusseault, Many remote learning options shutting down as school reopens for fall 2022. *The 74 Million* (2022).
2. A Zimmerman, 2023, p EDT, NYC public schools go remote Friday due to pollution from wildfire smoke (https://www.chalkbeat.org/newyork/2023/6/8/23754155/nyc-air-pollution-canada-wildfire-school-closure-remote-learning-friday/) (2023).
3. The School District of Philadelphia, A Message Regarding Election Day 2022 - The School District of Philadelphia (2022).
4. Digital Learning Collaborative, Snapshot 2024: The post pandemic digital learning landscape emerges, Technical report (2024).
5. S Cohodes, et al., Student achievement gaps and the pandemic: A new review of evidence from 2021–2022, (Center on reinventing public education), Technical report (2022).
6. The Nation's Report Card, NAEP Mathematics: Mathematics Highlights 2022 (https://www.nationsreportcard.gov/highlights/mathematics/2022/) (2022).
7. M Kuhfeld, J Soland, K Lewis, Test Score Patterns Across Three COVID-19-Impacted School Years. *Educ. Res.* **51**, 500–506 (2022).
8. R Jack, C Halloran, J Okun, E Oster, Pandemic Schooling Mode and Student Test Scores: Evidence from US School Districts. *Am. Econ. Rev. Insights* **5**, 173–190 (2023).
9. D Goldhaber, et al., The Educational Consequences of Remote and Hybrid Instruction during the Pandemic. *Am. Econ. Rev. Insights* **5**, 377–392 (2023).
10. AL Duckworth, PD Quinn, E Tsukayama, What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *J. Educ. Psychol.* **104**, 439–451 (2012).
11. LH Cross, RB Frary, Hodgepodge Grading: Endorsed by Students and Teachers Alike. *Appl. Meas. Educ.* **12**, 53–72 (1999).
12. WW Willingham, JM Pollack, C Lewis, Grades and Test Scores: Accounting for Observed Differences. *J. Educ. Meas.* **39**, 1–37 (2002).
13. BM Galla, et al., Why high school grades are better predictors of on-time college graduation than are admissions test scores: The roles of self-regulation and cognitive ability. *Am. Educ. Res. J.* **56**, 2077–2115 (2019).
14. CK Jackson, What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes. *J. Polit. Econ.* **126**, 2072–2107 (2018).
15. T Kautz, W Zanoni, Measurement and Development of Non-Cognitive Skills in Adolescence: Evidence from Chicago Public Schools and the OneGoal Program. *J. Hum. Cap.* **0**, 1–119 (2023).
16. R Chetty, D Deming, J Friedman, Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges, (National Bureau of Economic Research, Cambridge, MA), Technical Report w31492 (2023).
17. MA Kraft, Interpreting Effect Sizes of Education Interventions. *Educ. Res.* **49**, 241–253 (2020).
18. D Goldstein, What the New, Low Test Scores for 13-Year-Olds Say About U.S. Education Now. *The New York Times* (2023).
19. T Kane, S Reardon, Far Behind Their Kids Are in School. *The New York Times* (2023).
20. D Goldhaber, et al., The consequences of remote and hybrid instruction during the pandemic, (Center for Education Policy Research, Harvard University, Cambridge, MA), Technical report (2022).
21. E Sanchez, R Moore, Grade Inflation Continues to Grow in the Past Decade (2022).
22. D Arkhangelsky, S Athey, DA Hirshberg, GW Imbens, S Wager, Synthetic Difference-in-Differences. *Am. Econ. Rev.* **111**, 4088–4118 (2021).
23. AL Duckworth, JL Taxer, L Eskreis-Winkler, BM Galla, JJ Gross, Self-control and academic achievement. *Annu. Rev. Psychol.* **70**, 373–399 (2019).
24. L Borghans, BHH Golsteyn, JJ Heckman, JE Humphries, What grades and achievement tests measure. *Proc. Natl. Acad. Sci.* **113**, 13354–13359 (2016).
25. DJ Simons, Y Shoda, DS Lindsay, Constraints on Generality (COG): A Proposed Addition to All Empirical Papers. *Perspectives on Psychol. Sci.* **12**, 1123–1128 (2017).
26. AL Duckworth, et al., Students Attending School Remotely Suffer Socially, Emotionally, and Academically. *Educ. Res.* **50**, 479–482 (2021).

Supplementary Online Materials for

Declines in GPA During Remote Schooling Are Steepest for the Most Vulnerable Students

A. Methods

A. Participants and Procedure. All data for this investigation were collected from Orange County Public Schools in Florida by Character Lab Research Network (CLRN). We obtained students' course grades from school records and calculated GPAs on a 0-100 scale. GPAs were winsorized such that any value greater than 100 was coded as 100. GPAs were calculated by averaging the grades from the up to eight classes that students could take in a marking period.

Our main sample comprised 9,912 students in grades nine, ten, and eleven in the 2020-21 school year. The district has offered a full-time virtual academy option for students since before the COVID-19 pandemic. A trivial number of sample students enrolled in the virtual academy at baseline, indicating that inclusion or removal of these students from the study would make no difference in findings. Additionally, because students could change schools over time and enrollment at baseline did not necessarily indicate location of enrollment during subsequent marking periods, students ever enrolled in the virtual academy during the study period were retained in all analyses.

As a robustness check, we repeated analyses with a larger sample (N = 20,951) that also included sixth- and seventh-grade students, for whom the only demographic variables available in data received from the district were grade, gender, race/ethnicity, and baseline GPA. The 2019-20 marking period 1 GPA data were not available for these younger students, so the robustness check controls estimate the difference between remote and in-person students' GPA from times 2 to 12.

B. Analytic Strategy. Unless otherwise noted, all models are fixed effects panel models as explained in the main text. Specifically, we estimated the following model

$$y_{ist} = \alpha_i + \sum_{t=1}^{12} \psi_t + \left(\sum_{t=1}^{12} \lambda_t \times Remote_{is} \right) + \epsilon_{ist},$$

where y_{ist} is the GPA of student i in school s during time period t; α_i represents a student fixed effect; ψ_t represents changes in each time point for the group of students who did not attend remote schooling; λ_t represents our parameters of interest, the difference in GPA between in-person students and students who were remote during the 2020-21 academic year; $Remote_{is}$ is an indicator for whether a student attended school in person during the 2020-21 academic year; and ϵ_{ist} is an error term. To examine moderation effects, we estimated the difference between remote versus in-person students separately within each demographic subgroup. We use time t=4 as the reference period, so the estimated differences are relative to the difference between remote and in-person students during the period before remote schooling was optional. Standard errors were clustered at the individual level and were calculated using bootstrapping (b = 500). Effect sizes were calculated by standardizing all student-quarter observations before estimating the models.

We defined declines as $D = \lambda_8 - \lambda_5$ the difference between the estimated difference at times 8 and 5 (beginning and ending of the 2020-21 academic year). Similarly, recoveries were defined as $R = \lambda_9 - \lambda_8$ the difference between the first marking period after all students returned to in-person classes, and the final marking period when some students were remote and some in person. Differences in declines and recoveries between two groups were calculated as $D_i - D_j$ and $R_i - R_j$, where i and j index different demographic subgroups (e.g., male and female students). We used bootstrapping (b = 500) to compute standard errors for all estimates presented below and permutations to estimate a null distribution used for computing p-values.

As a robustness check for our main analyses, detailed in section D, we used two alternative model specification. Specifically, at each time point of the 2020-21 and 2021-22 academic year, we used an ordinary least squares (OLS) model to estimate the difference in GPAs of students attending remotely versus in person in 2020-21 when accounting for baseline GPA, demographics, and fixed effects for each school:

$$y_{ist} = \alpha + \beta Remote_{is} + \gamma X_{is} + \nu_s + \epsilon_{ist},$$

where y_{ist} is the GPA of student i in school s during time period t; α represents an intercept; β is the estimated disadvantage of attending school remotely; $Remote_{is}$ is an indicator for whether a student attended school remotely during the 2020-21 academic school year; X_{is} is a vector of control covariates, including grade, gender, race/ethnicity, eligibility for free or reduced-price meals, English language learner status, special education status, home language, and baseline GPA for each term in the 2019-20 school year; ν_s is a school fixed effect; and ϵ_{ist} is an error term. Reported standard errors were corrected for heteroskedasticity with the Huber-White sandwich estimator.

The other approach, reported in **Table S12** and **Figure S1**, was a synthetic difference-in-differences approach (22). This approach weights control units and previous timepoints to ensure comparability between the treated (i.e., remote) students and the control (i.e., in-person) students. We estimated separate models for each period post-intervention (i.e., times 5 - 12).

Students completed a survey reporting their remote versus in-person status at three points during the 2020-21 academic year. Our main specification uses students' latest response, in the spring of the 2020-21 year. These surveys were also administered to a random subsample of students in the fall and the winter of the academic year. This allows us to test an underlying assumption of our analyses: that remote versus in-person enrollment was consistent across the 2020-21 school year. Supporting this assumption, 76% of the students surveyed in the fall and 87% of students surveyed in the winter reported the same learning modality as in the spring. Additionally, deviations from this assumption would, if anything, dilute the negative effects of remote schooling, making our estimates conservative. The full survey including the question used for remote schooling status is available at https://doi.org/10.6084/m9.figshare.25043738.v1.

B. Results

A. Baseline differences. Baseline differences did not suggest that students more vulnerable to academic underachievement were more likely to attend school remotely. Tables S1 and S2 show the preexisting differences in the demographic characteristics of students who attended school in person versus remotely in the 2020-21 school year. Table S1 reports results of bivariate chi-square and t-tests from which the estimates and p-values presented in Figure 2 are derived. Table S2 shows that entering each variable at a time, or all simultaneously in probit models, produces similar results.

To estimate the differences in the linear trends in overall GPA during the first year, we fit a panel model with student fixed effects and linear time interacted with remote status, as follows.

$$y_{ist} = \alpha + \beta Quarter_t + \lambda Remote_{is} * Quarter_t + \delta_i + \epsilon_{ist}$$

where y_{ist} is the GPA of student i in school s during time period t in the 2019-20 academic year (t indexes from 1 to 4); α represents an intercept; β represents the average increase in GPA per quarter for students who would be in person in the 2020-21 academic year, λ represents the difference in the average quarterly increase in GPA between students who would be remote in 2020-21 and students who would be in person; as before, $Remote_{is}$ is an indicator for whether a student attended school in person during the 2020-21 academic year; δ_i is a student fixed effect; and ϵ_{ist} is an error term. Results from this model indicated that students who would eventually be remote not only started with higher GPAs, but their GPAs were increasing through the year at a faster pace ($b = 0.60$, SE $= 0.04$, $p < .001$, 95% CI [0.51 - 0.68]).

B. Main specification. High school students attending school remotely saw monotonic decreases in their GPAs relative to their peers who remained in person, even after controlling for all general time trends and time-invariant student-level covariates. The total extent of the decline during the remote period was 3.70 GPA points ($d = -0.37$). Upon returning to school in person, this gap quickly narrowed, with remote students recovering 2.22 GPA points ($d = 0.22$). **Table S3** shows results from our main specification using student fixed effects.

C. Moderation analyses. Tables S4 to S11 present moderation results by gender, race/ethnicity, income status, English language learner status, special education student status, home language, grade level, and baseline GPA, respectively. We found consistent moderation such that students who were male (Table S4), Black or Hispanic (Table S5), came from low-income families (Table S6), or who had lower baseline GPAs (Table S11) were more likely to see a disproportionately large impact of remote schooling. They saw the largest declines but also recovered faster. There were no consistent differences by English language learner status (Table S7), special education student status (Table S8), home language (Table S9), or grade level (Table S10).

D. Robustness checks. Table S12 shows sensitivity analyses. Specifically, we first analyzed the effect of remote schooling on different outcomes—Core GPA, Math GPA, and English Language Arts (ELA) GPA. Second, we tested robustness on different samples—a larger sample of data that included middle schoolers, the full sample that included middle and high school students, and the subset of students who reported the same learning location across all three surveys. Finally, we show that using an alternative OLS specification with demographic controls and school fixed effects, and which controls for the first four periods of GPA, produces similar results.

Table S1. Baseline Differences Between Students Who Chose Remote vs. In-Person Learning

Characteristic at Time 1	Mean (SE)		Δ (SE)	*p*
	Remote	In person		
Female	60.40% (0.73)	49.72% (0.68)	10.68% (1.00)	<.001
Race/ethnicity				<.001
Hispanic	41.32% (0.74)	38.94% (0.66)	2.38% (0.99)	.016
Black, non-Hispanic	22.66% (0.63)	21.85% (0.56)	0.81% (0.84)	.335
Asian, non-Hispanic	10.16% (0.45)	3.84% (0.26)	6.32% (0.52)	<.001
White, non-Hispanic	22.93% (0.63)	32.91% (0.64)	−9.97% (0.90)	<.001
Other race, non-Hispanic	2.93% (0.25)	2.47% (0.21)	0.46% (0.33)	.160
Free or reduced-price meal program	41.50% (0.74)	40.29% (0.66)	1.21% (0.99)	.224
English language learner	7.37% (0.39)	9.88% (0.40)	−2.52% (0.56)	<.001
Special education	26.67% (0.66)	29.42% (0.62)	−2.74% (0.91)	.002
Home language				<.001
English	58.37% (0.74)	63.95% (0.65)	−5.58% (0.98)	<.001
Spanish	26.00% (0.66)	25.00% (0.59)	1.00% (0.88)	.256
Other language	15.63% (0.55)	11.05% (0.42)	4.58% (0.69)	<.001
Grade level in 2020-21				<.001
9	33.54% (0.71)	42.04% (0.67)	−8.50% (0.97)	<.001
10	36.86% (0.72)	34.19% (0.64)	2.67% (0.97)	.006
11	29.60% (0.69)	23.77% (0.58)	5.83% (0.89)	<.001
Overall GPA at Time 1	87.02 (0.12)	84.84 (0.12)	2.17 (0.17)	<.001
Overall GPA at Time 2	86.08 (0.12)	83.64 (0.12)	2.44 (0.17)	<.001
Overall GPA at Time 3	86.09 (0.12)	83.25 (0.13)	2.84 (0.18)	<.001
Overall GPA at Time 4	89.18 (0.14)	85.15 (0.15)	4.02 (0.20)	<.001
Sample size	4,439	5,473		

Note. Δ is the difference between the Remote − In-person mean or proportion; positive values indicate that this demographic was more likely to attend school remotely. For continuous and dichotomous variables, the *p*-values were calculated using two-tailed t-tests. For categorical variables, the *p*-values were calculated using chi-square tests. * two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$. GPA = grade point average (100-point scale). Time 1 is Q1 of 2019-20.

Table S2. Marginal Effects of Baseline Characteristics on the Percent Chance of Choosing Remote Learning During 2020-21 (Main Analytic Sample)

Characteristic at Time 1	Unadjusted		Adjusted	
	Marginal effect	*p*	Marginal effect	*p*
Female (versus male)	10.65%***	<.001	6.45%***	<.001
Race/ethnicity	–	<.001	–	<.001
White, non-Hispanic (reference)	–	–	–	–
Hispanic	10.14%***	<.001	12.86%***	<.001
Black, non-Hispanic	9.57%***	<.001	12.32%***	<.001
Asian, non-Hispanic	32.12%***	<.001	26.14%***	<.001
Other race, non-Hispanic	12.94%***	<.001	12.25%***	<.001
Free or reduced-price meal program	1.24%	0.224	2.77%**	.009
English language learner	−7.79%***	<.001	−11.78%***	<.001
Special education	−3.35%**	0.002	−3.44%**	.002
Home language	–	<.001	–	<.001
English (reference)	–	–	–	–
Spanish	3.22%**	0.006	2.52%	.110
Other language	10.89%***	<.001	6.32%***	<.001
Grade level in 2020-21	–	<.001	–	<.001
9 (reference)	–	–	–	–
10	7.36%***	<.001	7.12%***	<.001
11	10.96%***	<.001	11.59%***	<.001
Overall GPA at Time 1 (Q1 2019-20)	0.77%***	<.001	−0.11%	.357
Overall GPA at Time 2 (Q2 2019-20)	0.80%***	<.001	−0.11%	.388
Overall GPA at Time 3 (Q3 2019-20)	0.90%***	<.001	0.38%***	<.001
Overall GPA at Time 4 (Q4 2019-20)	0.96%***	<.001	0.86%***	<.001
Sample size	9,912		9,912	

Note. The results are based on a probit model. The marginal effects were multiplied by 100, so they represent the effect of each covariate on the percent chance of choosing remote schooling. The *p*-values that appear next to individual variables come from chi-square tests of the null hypothesis that the effect is zero. The *p*-values that appear next to categorical variables come from chi-square tests of the null hypothesis that the effects are equal across categories.

* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S3. Main Specification

Term	Means: In person	Means: Remote	Adjusted difference: Standardized	Adjusted difference: Original units
Time 1	84.84	87.02	−0.19***	−1.85***
			(0.02)	(0.15)
Time 2	83.64	86.08	−0.16***	−1.58***
			(0.02)	(0.15)
Time 3	83.25	86.09	−0.12***	−1.18***
			(0.01)	(0.14)
Time 4	89.18	85.15		
Time 5	82.40	87.33	0.09***	0.90***
			(0.02)	(0.17)
Time 6	81.98	86.08	0.01	0.07
			(0.02)	(0.18)
Time 7	83.76	85.89	−0.19***	−1.90***
			(0.02)	(0.17)
Time 8	84.39	85.61	−0.28***	−2.80***
			(0.02)	(0.19)
Time 9	85.78	89.23	−0.06*	−0.58*
			(0.02)	(0.18)
Time 10	84.15	87.98	−0.02	−0.20
			(0.02)	(0.18)
Time 11	83.32	87.08	−0.03	−0.26
			(0.02)	(0.19)
Time 12	83.61	87.40	−0.02	−0.24
			(0.02)	(0.19)
Decline			−0.37***	−3.70***
			(0.02)	(0.17)
Recovery			0.22***	2.22***
			(0.02)	(0.17)
Sample Size	5,473	4,439		

Note. The models control for student and time fixed effects. Values in parentheses are bootstrapped standard errors (b = 500).
* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S4. Moderation Analyses for Gender

	Remote means		In-person means		Adjusted difference		Moderation
Term	Male	Female	Male	Female	Male	Female	Gender difference
Time 1	86.20	87.56	83.28	86.43	−1.70***	−1.73***	−0.03
					(0.24)	(0.19)	(0.30)
Time 2	85.01	86.78	81.83	85.46	−1.44***	−1.54***	−0.10
					(0.25)	(0.18)	(0.30)
Time 3	85.08	86.74	81.37	85.14	−0.91*	−1.25**	−0.35
					(0.22)	(0.18)	(0.28)
Time 4	87.57	90.23	82.96	87.37			
Time 5	85.92	88.25	80.44	84.38	0.86*	1.01**	0.15
					(0.26)	(0.20)	(0.34)
Time 6	84.68	87.00	80.37	83.61	−0.31	0.53	0.84
					(0.27)	(0.23)	(0.35)
Time 7	84.64	86.70	82.34	85.20	−2.32***	−1.35***	0.97
					(0.26)	(0.21)	(0.33)
Time 8	84.26	86.50	83.07	85.72	−3.43***	−2.07***	1.35*
					(0.29)	(0.23)	(0.37)
Time 9	88.54	89.67	84.21	87.38	−0.28	−0.56	−0.27
					(0.27)	(0.22)	(0.36)
Time 10	87.22	88.47	82.56	85.77	0.04	−0.15	−0.19
					(0.27)	(0.22)	(0.36)
Time 11	86.03	87.77	81.68	84.99	−0.27	−0.07	0.20
					(0.29)	(0.23)	(0.37)
Time 12	86.40	88.05	82.04	85.20	−0.26	−0.01	0.25
					(0.30)	(0.24)	(0.38)
Decline					−4.29***	−3.08***	1.20*
					(0.26)	(0.20)	(0.33)
Recovery					3.14***	1.52***	−1.63**
					(0.27)	(0.22)	(0.35)
Sample size	2,752	2,721	1,758	2,681			

Note. The models control for student and time fixed effects. Values in parentheses are bootstrapped standard errors (b = 500).
* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S5. Moderation Analyses for Race/Ethnicity

Term	Remote means			In-person means			Adjusted difference			Moderation		
	Black/Hispanic	White/Other	Asian	Black/Hispanic	White/Other	Asian	Black/Hispanic	White/Other	Asian	WO - BH	A - WO	A - BH
Time 1	85.19	89.42	92.38	82.61	88.11	90.15	−2.05***	−1.58***	−1.51	0.47	0.08	0.54
							(0.21)	(0.24)	(0.58)	(0.32)	(0.65)	(0.61)
Time 2	84.48	88.07	91.03	81.62	86.53	88.86	−1.77***	−1.35*	−1.57	0.42	−0.21	0.20
							(0.21)	(0.24)	(0.58)	(0.31)	(0.64)	(0.60)
Time 3	84.56	87.95	90.93	81.27	86.06	88.60	−1.34***	−1.00*	−1.42	0.34	−0.41	−0.07
							(0.20)	(0.23)	(0.55)	(0.30)	(0.62)	(0.58)
Time 4	87.61	91.27	93.71	82.97	88.37	89.97						
Time 5	85.67	89.48	92.27	79.78	86.30	87.92	1.26***	0.27	0.61	−0.99	0.34	−0.65
							(0.22)	(0.27)	(0.64)	(0.36)	(0.71)	(0.67)
Time 6	84.29	88.21	91.92	79.43	85.92	86.06	0.22	−0.60	2.12	−0.82	2.73*	1.91
							(0.24)	(0.30)	(0.70)	(0.38)	(0.76)	(0.74)
Time 7	84.09	88.07	91.61	81.77	86.76	87.78	−2.31***	−1.59**	0.09	0.72	1.68	2.40*
							(0.24)	(0.29)	(0.70)	(0.38)	(0.78)	(0.76)
Time 8	83.77	87.76	91.71	82.68	86.85	88.62	−3.54***	−1.99***	−0.64	1.56*	1.34	2.90**
							(0.25)	(0.31)	(0.69)	(0.39)	(0.77)	(0.74)
Time 9	87.91	90.97	93.07	83.91	88.48	90.55	−0.64	−0.41	−1.22	0.23	−0.81	−0.58
							(0.23)	(0.29)	(0.63)	(0.37)	(0.70)	(0.67)
Time 10	86.53	89.87	92.24	82.21	86.95	89.05	−0.32	0.02	−0.55	0.34	−0.57	−0.24
							(0.23)	(0.30)	(0.66)	(0.40)	(0.75)	(0.69)
Time 11	85.49	89.25	91.57	81.18	86.46	88.32	−0.33	−0.10	−0.49	0.22	−0.38	−0.16
							(0.24)	(0.31)	(0.67)	(0.39)	(0.78)	(0.71)
Time 12	85.86	89.47	91.79	81.63	86.49	88.59	−0.40	0.08	−0.54	0.48	−0.62	−0.14
							(0.25)	(0.32)	(0.74)	(0.41)	(0.83)	(0.79)
Decline							−4.80***	−2.26***	−1.26	2.54***	1.00	3.55**
							(0.22)	(0.27)	(0.57)	(0.35)	(0.63)	(0.63)
Recovery							2.91***	1.58***	−0.57	−1.33*	−2.15	−3.48***
							(0.23)	(0.30)	(0.54)	(0.38)	(0.63)	(0.59)
Sample size	3,327	1,936	210	2,840	1,148	451						

Note. The models control for student and time fixed effects. Values in parentheses are bootstrapped standard errors (b = 500). A = Asian, WO = White or Other race, BH = Black or Hispanic

* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S6. Moderation Analyses for Low-Income Status

	Remote means		In-person means		Adjusted difference		Moderation
Term	Not LI	LI	Not LI	LI	Not LI	LI	LI difference
Time 1	88.34	85.15	86.88	81.83	−1.34***	−2.62***	−1.28*
					(0.17)	(0.27)	(0.31)
Time 2	87.24	84.45	85.59	80.74	−1.15***	−2.23***	−1.07*
					(0.17)	(0.26)	(0.30)
Time 3	87.19	84.53	85.26	80.27	−0.87*	−1.67***	−0.80
					(0.17)	(0.25)	(0.30)
Time 4	90.53	87.27	87.72	81.34			
Time 5	88.70	85.39	84.98	78.58	0.92***	0.88	−0.03
					(0.19)	(0.29)	(0.36)
Time 6	87.49	84.08	84.49	78.27	0.20	−0.12	−0.32
					(0.20)	(0.30)	(0.36)
Time 7	87.29	83.90	85.83	80.70	−1.34***	−2.73***	−1.39**
					(0.21)	(0.30)	(0.36)
Time 8	87.10	83.52	86.31	81.53	−2.02***	−3.94***	−1.93***
					(0.21)	(0.32)	(0.39)
Time 9	90.36	87.62	87.61	83.08	−0.05	−1.39***	−1.34**
					(0.20)	(0.29)	(0.36)
Time 10	89.20	86.25	86.14	81.21	0.26	−0.89*	−1.15*
					(0.20)	(0.30)	(0.37)
Time 11	88.37	85.26	85.44	80.19	0.13	−0.86*	−0.99
					(0.21)	(0.30)	(0.37)
Time 12	88.63	85.66	85.72	80.49	0.10	−0.76	−0.86
					(0.23)	(0.32)	(0.39)
Decline					−2.93***	−4.83***	−1.90***
					(0.19)	(0.30)	(0.36)
Recovery					1.97***	2.56***	0.59
					(0.20)	(0.30)	(0.35)
Sample size	3,268	2,105	2,597	1,842			

Note. The models control for student and time fixed effects. Values in parentheses are bootstrapped standard errors (b = 500). LI = low income.
* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S7. Moderation Analyses for English Language Learner Status

	Remote means		In-person means		Adjusted difference		Moderation
Term	Not ELL	ELL	Not ELL	ELL	Not ELL	ELL	ELL difference
Time 1	87.36	82.75	85.32	80.55	−1.90***	−2.09*	−0.19
					(0.16)	(0.59)	(0.61)
Time 2	86.31	83.12	83.97	80.62	−1.59***	−1.79	−0.19
					(0.16)	(0.54)	(0.55)
Time 3	86.25	84.07	83.50	80.92	−1.19***	−1.14	0.06
					(0.15)	(0.54)	(0.54)
Time 4	89.32	87.36	85.38	83.07			
Time 5	87.50	85.10	82.71	79.56	0.85***	1.26	0.40
					(0.16)	(0.58)	(0.60)
Time 6	86.20	84.48	82.19	80.08	0.07	0.11	0.04
					(0.19)	(0.60)	(0.63)
Time 7	86.00	84.42	83.99	81.73	−1.93***	−1.59	0.34
					(0.18)	(0.61)	(0.64)
Time 8	85.68	84.80	84.56	82.77	−2.83***	−2.25*	0.58
					(0.19)	(0.64)	(0.67)
Time 9	89.41	86.91	86.05	83.32	−0.58	−0.69	−0.11
					(0.18)	(0.61)	(0.64)
Time 10	88.14	85.94	84.39	81.98	−0.19	−0.32	−0.13
					(0.18)	(0.62)	(0.66)
Time 11	87.26	84.88	83.62	80.62	−0.30	−0.03	0.28
					(0.18)	(0.64)	(0.67)
Time 12	87.52	85.78	83.85	81.48	−0.26	0.01	0.28
					(0.19)	(0.66)	(0.69)
Decline					−3.68***	−3.51***	0.17
					(0.17)	(0.60)	(0.63)
Recovery					2.24***	1.56	−0.69
					(0.18)	(0.61)	(0.64)
Sample size	4,932	541	4,112	327			

Note. The models control for student and time fixed effects. Values in parentheses are bootstrapped standard errors (b = 500). ELL = English language learner
* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S8. Moderation Analyses for Special Education Status

	Remote means		In-person means		Adjusted difference		Moderation
Term	Not SPED	SPED	Not SPED	SPED	Not SPED	SPED	SPED difference
Time 1	86.41	88.70	83.92	87.07	−1.88***	−1.74*	0.14
					(0.18)	(0.27)	(0.32)
Time 2	85.53	87.59	82.78	85.70	−1.62***	−1.48**	0.14
					(0.18)	(0.25)	(0.31)
Time 3	85.53	87.62	82.35	85.39	−1.19***	−1.14*	0.05
					(0.17)	(0.25)	(0.30)
Time 4	88.67	90.57	84.30	87.20			
Time 5	86.83	88.69	81.47	84.64	1.00**	0.68	−0.32
					(0.19)	(0.27)	(0.34)
Time 6	85.55	87.52	80.95	84.46	0.24	−0.32	−0.55
					(0.21)	(0.30)	(0.35)
Time 7	85.26	87.59	82.89	85.86	−2.00***	−1.64***	0.35
					(0.21)	(0.29)	(0.36)
Time 8	84.99	87.32	83.52	86.45	−2.90***	−2.50***	0.40
					(0.21)	(0.32)	(0.38)
Time 9	88.73	90.58	85.09	87.44	−0.73*	−0.23	0.49
					(0.20)	(0.31)	(0.37)
Time 10	87.42	89.50	83.34	86.10	−0.29	0.03	0.32
					(0.20)	(0.31)	(0.37)
Time 11	86.49	88.72	82.44	85.45	−0.32	−0.10	0.22
					(0.22)	(0.32)	(0.39)
Time 12	86.82	88.97	82.74	85.72	−0.28	−0.12	0.16
					(0.22)	(0.33)	(0.40)
Decline					−3.90***	−3.18***	0.72
					(0.20)	(0.29)	(0.35)
Recovery					2.17***	2.27***	0.10
					(0.21)	(0.30)	(0.38)
Sample size	3,863	1,610	3,255	1,184			

Note. The models control for student and time fixed effects. Values in parentheses are bootstrapped standard errors (b = 500). SPED = special education student.
* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S9. Moderation Analyses for Home Language

	Remote means			In-person means			Adjusted difference			Moderation		
Term	English	Spanish	Other	English	Spanish	Other	English	Spanish	Other	S-O	E-S	E-O
Time 1	87.47	85.15	88.43	85.70	82.80	84.54	−1.73***	−2.45***	−1.18	1.27	−0.72	0.55
							(0.19)	(0.32)	(0.43)	(0.53)	(0.36)	(0.46)
Time 2	86.43	84.51	87.36	84.28	81.88	83.88	−1.36***	−2.17***	−1.59	0.58	−0.81	−0.23
							(0.18)	(0.31)	(0.42)	(0.51)	(0.35)	(0.45)
Time 3	86.32	84.68	87.56	83.75	81.77	83.67	−0.94***	−1.89***	−1.18	0.71	−0.95	−0.24
							(0.18)	(0.30)	(0.39)	(0.48)	(0.35)	(0.42)
Time 4	89.47	87.87	90.26	85.96	83.06	85.18						
Time 5	87.66	85.58	88.99	83.34	80.06	82.26	0.82*	0.71	1.65*	0.94	−0.10	0.84
							(0.21)	(0.37)	(0.43)	(0.57)	(0.44)	(0.49)
Time 6	86.37	84.32	87.90	82.88	79.72	81.89	−0.01	−0.20	0.93	1.13	−0.19	0.95
							(0.22)	(0.38)	(0.48)	(0.62)	(0.44)	(0.53)
Time 7	86.18	84.01	87.92	84.57	81.62	83.97	−1.90***	−2.41***	−1.13	1.28	−0.52	0.77
							(0.22)	(0.37)	(0.49)	(0.63)	(0.44)	(0.54)
Time 8	85.77	83.89	87.90	84.92	82.82	84.82	−2.66***	−3.73***	−1.99**	1.74	−1.07	0.67
							(0.22)	(0.40)	(0.49)	(0.63)	(0.46)	(0.54)
Time 9	89.59	87.77	90.28	86.43	84.14	85.77	−0.34	−1.17*	−0.57	0.60	−0.82	−0.22
							(0.21)	(0.37)	(0.48)	(0.62)	(0.43)	(0.53)
Time 10	88.36	86.21	89.50	84.82	82.46	84.15	0.03	−1.05	0.27	1.33	−1.09	0.24
							(0.21)	(0.38)	(0.50)	(0.64)	(0.44)	(0.55)
Time 11	87.52	85.26	88.47	84.07	81.39	83.35	−0.06	−0.94	0.05	0.98	−0.88	0.10
							(0.23)	(0.40)	(0.52)	(0.68)	(0.47)	(0.57)
Time 12	87.78	85.63	88.90	84.28	81.90	83.62	−0.01	−1.08*	0.21	1.28	−1.07	0.21
							(0.24)	(0.40)	(0.52)	(0.69)	(0.48)	(0.58)
Decline							−3.48***	−4.45***	−3.64***	0.80	−0.97	−0.17
							(0.20)	(0.35)	(0.46)	(0.57)	(0.40)	(0.50)
Recovery							2.32***	2.57***	1.42	−1.14	0.25	−0.90
							(0.20)	(0.35)	(0.45)	(0.57)	(0.39)	(0.49)
Sample size	3,500	1,368	605	2,591	1,154	694						

Note. The models control for student and time fixed effects. Values in parentheses are bootstrapped standard errors (b = 500). E = English, S = Spanish, O = Other. The last three columns represent pairwise differences for each pair of home languages.
* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S10. Moderation Analyses for Grade Level

Term	Remote means			In-person means			Adjusted difference			Moderation		
	8th Grade	9th Grade	10th Grade	8th Grade	9th Grade	10th Grade	8th Grade	9th Grade	10th Grade	10 - 9	10 - 8	9 - 8
Time 1	87.85	86.99	86.10	86.11	84.13	83.64	−1.81***	−1.73***	−1.79***	−0.05	0.02	0.07
							(0.23)	(0.26)	(0.30)	(0.40)	(0.39)	(0.33)
Time 2	86.66	86.37	85.05	84.75	83.09	82.45	−1.64**	−1.32**	−1.64**	−0.32	0.01	0.33
							(0.24)	(0.26)	(0.29)	(0.38)	(0.39)	(0.34)
Time 3	86.56	86.26	85.34	84.30	82.87	81.92	−1.29**	−1.21**	−0.83	0.38	0.46	0.07
							(0.23)	(0.24)	(0.29)	(0.38)	(0.37)	(0.32)
Time 4	89.46	89.35	88.64	85.90	84.75	84.40						
Time 5	87.27	87.14	87.63	82.95	81.77	82.33	0.77	0.77	1.05*	0.28	0.28	0.00
							(0.28)	(0.26)	(0.32)	(0.43)	(0.45)	(0.38)
Time 6	86.10	86.04	86.10	82.98	81.18	81.37	−0.43	0.26	0.49	0.22	0.92	0.70
							(0.29)	(0.28)	(0.36)	(0.46)	(0.45)	(0.39)
Time 7	85.72	86.06	85.85	84.39	83.36	83.23	−2.22***	−1.90***	−1.62**	0.28	0.60	0.32
							(0.28)	(0.28)	(0.36)	(0.47)	(0.45)	(0.39)
Time 8	85.22	85.63	86.03	84.85	84.04	84.06	−3.18***	−3.01***	−2.28***	0.73	0.90	0.17
							(0.31)	(0.30)	(0.36)	(0.48)	(0.48)	(0.42)
Time 9	88.41	89.14	90.26	85.35	85.43	87.06	−0.50	−0.88*	−1.05	−0.16	−0.55	−0.39
							(0.27)	(0.28)	(0.36)	(0.46)	(0.46)	(0.37)
Time 10	87.01	88.17	88.83	83.65	83.85	85.47	−0.19	−0.28	−0.89	−0.61	−0.70	−0.09
							(0.28)	(0.28)	(0.36)	(0.47)	(0.47)	(0.39)
Time 11	86.12	87.23	87.99	82.83	83.01	84.64	−0.26	−0.37	−0.90	−0.52	−0.63	−0.11
							(0.30)	(0.29)	(0.38)	(0.48)	(0.49)	(0.41)
Time 12	86.26	87.67	88.34	83.01	83.15	85.34	−0.31	−0.08	−1.24*	−1.16	−0.94	0.23
							(0.32)	(0.30)	(0.38)	(0.50)	(0.50)	(0.43)
Decline							−3.94***	−3.78***	−3.33***	0.45	0.62	0.17
							(0.29)	(0.28)	(0.31)	(0.42)	(0.46)	(0.42)
Recovery							2.68***	2.13***	1.23*	−0.90	−1.45*	−0.56
							(0.28)	(0.28)	(0.32)	(0.41)	(0.43)	(0.40)
Sample size	2,301	1,871	1,301	1,489	1,636	1,314						

Note. The models control for student and time fixed effects. Values in parentheses are bootstrapped standard errors (b = 500).
* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S11. Moderation Analyses for Baseline GPA

Term	Remote means		In-person means		Adjusted difference		Moderation
	Above median	Below median	Above median	Below median	Above median	Below median	Baseline GPA difference
Time 1	92.05	79.98	91.99	79.30	−0.58	−2.58***	2.01***
					(0.12)	(0.29)	(0.32)
Time 2	91.35	78.71	91.17	77.78	−0.47	−2.34***	1.87***
					(0.12)	(0.29)	(0.31)
Time 3	91.30	78.79	91.06	77.18	−0.40	−1.65***	1.25*
					(0.11)	(0.27)	(0.29)
Time 4	94.62	81.57	93.97	78.30			
Time 5	92.41	80.22	90.72	75.94	1.04***	1.02*	0.03
					(0.15)	(0.30)	(0.34)
Time 6	91.59	78.37	89.96	75.79	0.99**	−0.68	1.67***
					(0.16)	(0.32)	(0.36)
Time 7	91.15	78.53	90.53	78.51	−0.03	−3.25***	3.22***
					(0.16)	(0.31)	(0.35)
Time 8	90.85	78.30	90.92	79.31	−0.72	−4.28***	3.55***
					(0.18)	(0.32)	(0.37)
Time 9	92.58	84.54	91.28	81.51	0.65	−0.24	0.89
					(0.16)	(0.30)	(0.34)
Time 10	91.70	82.77	90.12	79.52	0.93*	−0.01	0.95
					(0.17)	(0.30)	(0.36)
Time 11	91.18	81.35	89.67	78.39	0.87*	−0.31	1.17
					(0.18)	(0.32)	(0.38)
Time 12	91.35	81.87	89.77	78.83	0.93*	−0.22	1.16*
					(0.19)	(0.33)	(0.38)
Decline					−1.77***	−5.29***	3.53***
					(0.17)	(0.29)	(0.34)
Recovery					1.37***	4.04***	−2.66***
					(0.18)	(0.30)	(0.35)
Sample size	3,081	2,392	1,851	2,588			

Note. The models control for student and time fixed effects. Values in parentheses are bootstrapped standard errors (b = 500).
* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S12. Robustness Checks

Term	Main specification	Different outcome			Different sample			Different estimation		
	Overall GPA	Core GPA	Math GPA	ELA GPA	Overall GPA	Overall GPA	Overall GPA	Overall GPA	Overall GPA	Overall GPA
Time 1	−1.85***	−1.93***	−1.98***	−1.72***		−2.01***	−2.08***		−1.85***	
	(0.15)	(0.17)	(0.27)	(0.27)		(0.14)	(0.25)		(0.19)	
Time 2	−1.58***	−1.63***	−1.43***	−1.47***	−1.26***	−1.54***	−1.82***		−1.58***	
	(0.15)	(0.16)	(0.28)	(0.26)	(0.16)	(0.11)	(0.25)		(0.18)	
Time 3	−1.18***	−1.21***	−1.11*	−1.11***	−0.87***	−1.08***	−1.53***		−1.18***	
	(0.14)	(0.16)	(0.27)	(0.24)	(0.15)	(0.10)	(0.24)		(0.18)	
Time 5	0.90***	0.87**	0.90*	0.56	0.60*	0.77***	−0.05	1.76***	0.90***	1.81***
	(0.17)	(0.19)	(0.35)	(0.28)	(0.17)	(0.12)	(0.28)	(0.15)	(0.24)	(0.15)
Time 6	0.07	0.34	0.29	−0.18	−0.41	−0.14	−2.03***	1.08***	0.07	0.94***
	(0.18)	(0.20)	(0.33)	(0.29)	(0.18)	(0.12)	(0.26)	(0.17)	(0.24)	(0.16)
Time 7	−1.90***	−1.42***	−1.53***	−2.35***	−2.98***	−2.44***	−2.23***	−0.68***	−1.90***	−0.90***
	(0.17)	(0.20)	(0.32)	(0.30)	(0.17)	(0.12)	(0.27)	(0.16)	(0.29)	(0.16)
Time 8	−2.80***	−2.37***	−2.98***	−3.13***	−4.56***	−3.75***	−2.62***	−1.77***	−2.80***	−1.82***
	(0.19)	(0.21)	(0.34)	(0.32)	(0.18)	(0.12)	(0.30)	(0.17)	(0.29)	(0.17)
Time 9	−0.58*	−0.33	−0.58	−0.47	−0.36	−0.55***	−0.77	1.35***	−0.58*	0.74***
	(0.18)	(0.20)	(0.32)	(0.28)	(0.17)	(0.12)	(0.30)	(0.13)	(0.21)	(0.15)
Time 10	−0.20	−0.04	−0.39	−0.08	−0.08	−0.22	−0.34	1.52***	−0.20	1.05***
	(0.18)	(0.21)	(0.35)	(0.29)	(0.17)	(0.13)	(0.31)	(0.14)	(0.20)	(0.15)
Time 11	−0.26	−0.14	−0.46	−0.22	−0.26	−0.33	−0.46	1.32***	−0.26	0.94***
	(0.19)	(0.22)	(0.36)	(0.29)	(0.18)	(0.12)	(0.32)	(0.16)	(0.22)	(0.16)
Time 12	−0.24	−0.02	−0.08	−0.05	−0.08	−0.26	−0.32	1.34***	−0.24	0.94***
	(0.19)	(0.23)	(0.36)	(0.30)	(0.19)	(0.13)	(0.34)	(0.16)	(0.22)	(0.17)
Decline	−3.70***	−3.24***	−3.88***	−3.69***	−5.16***	−4.52***	−2.57***		−3.70***	
	(0.17)	(0.17)	(0.31)	(0.25)	(0.18)	(0.11)	(0.26)		(0.28)	
Recovery	2.22***	2.04***	2.40***	2.67***	4.20***	3.21***	1.86***		2.22***	
	(0.17)	(0.19)	(0.35)	(0.29)	(0.16)	(0.12)	(0.26)		(0.21)	
Sample	High school	High school	High school	High school	Middle school	All students	Same location	High school	High school	High school
Estimation	Panel	Panel	Panel	Panel	Panel	Panel	Panel	OLS	Panel	Synthetic DID
SE Clustering	Student	Student	Student	Student	Student	Student	Student	Student	School	Student
Sample size	9,912	9,781	8,333	8,692	10,914	20,951	3,098	9,912	9,912	9,912

Note. Panel models control for student and time fixed effects. OLS models control for Time 1 (Q1 2019-20) grade, gender, race/ethnicity, eligibility for free or reduced-price meals, English language learner status, special education status, home language, school, and Time 1 (Q1 2019-20) through Time 4 (Q4 2019-20) overall GPA. For OLS models, each estimate, with its corresponding standard error, comes from separate regressions (i.e., the dependent variable is the GPA at times 5 - 12). For panel models, values in parentheses are bootstrapped standard errors (b = 500). Standard errors in the synthetic difference-in-difference estimation come from placebo variance estimation (b = 400). Main specification reproduced from Table S3 for comparison purposes. The sample labeled "same location" is the subset of students in the main analytic sample who also took the Fall 2020, winter 2021, and spring 2021 Character Lab Thriving Index Surveys and reported the same learning location across all three time points. Clustering in bootstrapped standard errors was achieved by resampling at the student and school levels. ELA = English language arts.
* two-tailed $p < .05$. ** two-tailed $p < .01$. *** two-tailed $p < .001$.

Table S13. Means by Learning Location for Alternative Samples

Time	High school		Middle school		All students		Same location	
	In person	Remote	In-Person	Remote	In person	Remote	In person	Remote
Time 1	84.84	87.02			84.84	87.02	87.41	89.15
Time 2	83.64	86.08	83.35	86.32	83.45	86.16	86.19	88.20
Time 3	83.25	86.09	82.60	85.95	82.84	86.01	85.84	88.13
Time 4	85.15	89.18	84.25	88.47	84.61	88.85	87.81	91.64
Time 5	82.40	87.33	81.49	86.32	81.85	86.86	86.46	90.24
Time 6	81.98	86.08	81.10	84.92	81.46	85.56	87.56	89.36
Time 7	83.76	85.89	83.41	84.65	83.54	85.33	87.32	88.92
Time 8	84.39	85.61	84.89	84.55	84.65	85.14	87.30	88.50
Time 9	85.78	89.23	85.38	89.25	85.53	89.23	88.07	91.13
Time 10	84.15	87.98	83.79	87.93	83.92	87.94	86.53	90.01
Time 11	83.32	87.08	82.99	86.96	83.10	87.02	85.99	89.35
Time 12	83.61	87.40	83.40	87.55	83.46	87.45	86.11	89.62

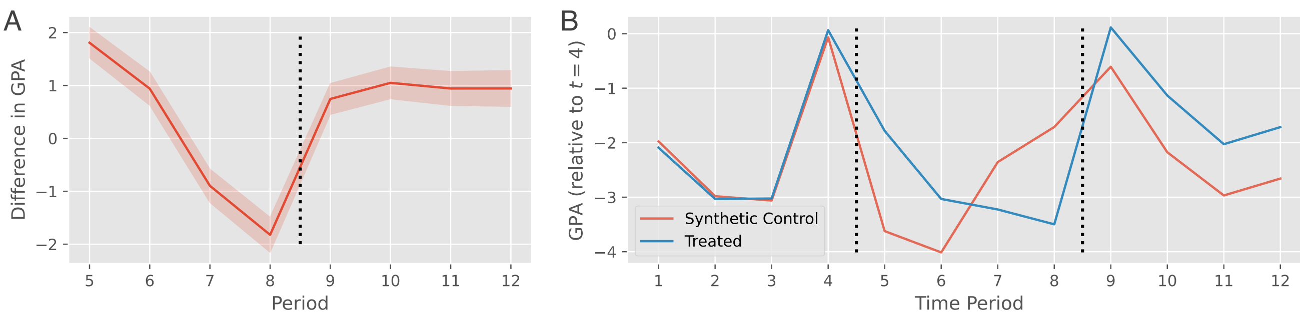

Fig. S1. Synthetic difference-in-differences results. (A) Effects at each time point with synthetic differences at each time point. (B) Means at each time point relative to $t = 4$ for remote students and a synthetic control version of in-person students.

Table S14. Means by Learning Location for Alternative Outcomes

Time	Overall GPA		Core GPA		Math GPA		ELA GPA	
	In person	Remote	In person	Remote	In person	Remote	In person	Remote
Time 1	84.84	87.02	82.28	84.65	81.01	82.82	82.80	85.64
Time 2	83.64	86.08	82.10	84.77	80.62	82.97	82.39	85.47
Time 3	83.25	86.09	81.65	84.73	79.97	82.66	82.36	85.80
Time 4	85.15	89.18	84.09	88.38	83.15	86.95	84.02	88.58
Time 5	82.40	87.33	80.65	85.81	78.59	83.28	82.40	87.51
Time 6	81.98	86.08	80.09	84.72	78.94	83.03	81.33	85.70
Time 7	83.76	85.89	81.59	84.45	79.63	81.90	83.48	85.69
Time 8	84.39	85.61	82.36	84.28	80.78	81.60	83.63	85.06
Time 9	85.78	89.23	83.24	87.18	80.61	83.83	84.89	88.99
Time 10	84.15	87.98	81.83	86.06	80.05	83.46	83.20	87.68
Time 11	83.32	87.08	80.98	85.11	78.47	81.80	82.61	86.95
Time 12	83.61	87.40	81.38	85.62	79.00	82.71	82.25	86.76