Online Appendix for:

School Closures and Effective In-Person Learning during COVID-19

A EIPL estimation approach

Our main methodological contribution is to derive and estimate a measure of effective in-person learning (EIPL) from the information contained in school visits and schooling mode trackers. This section provides details on the derivation and estimation procedure.

Derivation. Consider school j in county c in week t.[1] For a given number of in-person learning days $d_{j,t}$, a tracker categorizes the instructional mode of this school as one of three mutually exclusive cases:

- Traditional mode: $T_{j,t} = 100$ if $d_{j,t} \geq \bar{d}$
- Hybrid mode: $H_{j,t} = 100$ if $\underline{d} < d_{j,t} < \bar{d}$
- Virtual mode: $V_{j,t} = 100$ if $d_{j,t} \leq \underline{d}$

where $\underline{d}$ and $\bar{d}$ are the lower and upper thresholds for hybrid mode and vary by tracker. The three modes relate to effective in-person learning (EIPL) as follows:

$$EIPL_{j,t} = \begin{cases} 100 - \mu_j^T + \eta_{j,t}^T & \text{if } d_{j,t} \geq \bar{d} \\ 100\gamma_j + \eta_{j,t}^H & \text{if } \underline{d} < d_{j,t} < \bar{d} \\ \mu_j^V + \eta_{j,t}^V & \text{if } d_{j,t} \leq \underline{d} \end{cases} \tag{A.1}$$

where $\mu_j^T \geq 0$ represents the average deviation from 100% in-person learning when the tracker is in traditional mode; γ_j represents the average share of in-person learning when the tracker is in hybrid mode; $\mu_j^V \geq 0$ represents the average deviation from 0% in-person learning when the tracker is in virtual mode; and $\eta_{j,t}^T$, $\eta_{j,t}^H$, $\eta_{j,t}^V$ are mean-zero error terms. Student-weighted average EIPL for county c can therefore be expressed as:

$$\begin{aligned} EIPL_{c,t} &= \sum_{j \in c} \omega_j \left[\left(100 - \mu_j^T + \eta_{j,t}^T\right) \mathbb{1}\left\{d_{j,t} \geq \bar{d}\right\} \right. \\ &\qquad \left. + \left(100\gamma_j + \eta_{j,t}^H\right) \mathbb{1}\left\{\underline{d} < d_{j,t} < \bar{d}\right\} + \left(\mu_j^V + \eta_{j,t}^V\right) \mathbb{1}\left\{d_{j,t} \leq \underline{d}\right\} \right] \\ &= \left(100 - \tilde{\mu}_c^T + \tilde{\eta}_{c,t}^T\right) P_{c,t}^T + \left(100\gamma_c + \tilde{\eta}_{c,t}^H\right) P_{c,t}^H + \left(\tilde{\mu}_c^V + \tilde{\eta}_{c,t}^V\right) P_{c,t}^V \\ &= T_{c,t} + \gamma_c H_{c,t} - \left(\mu_c^T - \mu_c^V\right) + \left(\eta_{c,t}^T + \eta_{c,t}^H + \eta_{c,t}^V\right), \end{aligned} \tag{A.2}$$

where in the second line, $\tilde{\mu}_c^T$, $\tilde{\mu}_c^V$, $\tilde{\eta}_{c,t}^T$, $\tilde{\eta}_{c,t}^H$ and $\tilde{\eta}_{c,t}^V$ are student-weighted county averages of the above-defined deviations and error terms; γ_c is the student-weighted county average share of in-person learning when the tracker is in hybrid mode; and $P_{c,t}^T$, $P_{c,t}^H$, $P_{c,t}^V$ denote the share of schools in county c that are in each of the three modes. In the last line, $T_{c,t} = 100P_{c,t}^T$ and $H_{c,t} = 100P_{c,t}^H$ are the percentages of students

[1] Most trackers report instructional mode at either the school district or county level (see Subsection B.1). Some trackers build their measure up from information about individual schools (e.g. R2L), or weigh by instructional modes of different grades within districts (e.g. Burbio). For illustration, we assume here that the tracker measures instructional mode at the school level and reports data at the county level. All derivations go through if we assumed grade-level measures of instructional mode or district-level reporting.

in traditional and hybrid mode, respectively, that the tracker reports at the county level; and μ_c^T, μ_c^V, $\eta_{c,t}^T$, $\eta_{c,t}^H$, $\eta_{c,t}^V$ are the deviations and error terms weighted by the shares $P_{c,t}^T$, $P_{c,t}^H$ and $P_{c,t}^V$.

Equation (A.2) illustrates the two problems of trackers in measuring EIPL: (i) the proportion of in-person learning γ_c that a tracker's definition of hybrid mode implies may vary across regions; and (ii) trackers may attribute different learning mode percentages to different regions and, consequently, measure EIPL with more or less error (as captured by μ_c^T, γ_c, μ_c^V, $\eta_{c,t}^T$, $\eta_{c,t}^H$, $\eta_{c,t}^V$).

Our approach is based on the idea that since $EIPL_{c,t}$ and Safegraph visits $\Delta\widetilde{v}_{c,t}$ are in percent deviations from the pre-pandemic baseline of 100, the relationship between the two variables should satisfy $EIPL_{c,t} = 100 + \beta_c \Delta\widetilde{v}_{c,t} + \varepsilon_{c,t}$, where $\varepsilon_{c,t}$ is the measurement error implied by the Safegraph data Replacing with the above definition of $EIPL_{c,t}$, we therefore have the following linear regression

$$T_{c,t} = (100 + \mu_c) + \beta_c \Delta\widetilde{v}_{c,t} - \gamma_c H_{c,t} + (\varepsilon_{c,t} - \eta_{c,t}), \tag{A.3}$$

where $\mu_c = \mu_c^T - \mu_c^V$ and $\eta_{c,t} = \eta_{c,t}^T + \eta_{c,t}^H + \eta_{c,t}^V$. This is equation (4) in the main text.

Estimation. As explained in the main text, we estimate (A.3) separately at the Core-Based Statistical Areas (CBSA) and state level with either Burbio and R2L data.[2] Specifically, we implement the following three steps:

1. Restrict the data to counties or school districts for which our Safegraph-NCES dataset includes at least 5 schools.
2. For each reagion (CBSA, state) and each school tracker (Burbio, R2L), estimate equation (A.3) for different sample periods of the 2020-21 school year and retain the estimate with the highest R-squared.[3] This yields four pairs of estimates $\left\{\hat{\beta}_{\text{Burbio-CBSA}}, \hat{\gamma}_{\text{Burbio-CBSA}}\right\}$, $\left\{\hat{\beta}_{\text{R2L-CBSA}}, \hat{\gamma}_{\text{R2L-CBSA}}\right\}$, $\left\{\hat{\beta}_{\text{Burbio-state}}, \hat{\gamma}_{\text{Burbio-state}}\right\}$, and $\left\{\hat{\beta}_{\text{R2L-state}}, \hat{\gamma}_{\text{R2L-state}}\right\}$.
3. Rank the estimates based on the associated R-squared, denoted by $\left\{\hat{\beta}_{(k)}, \hat{\gamma}_{(\text{k})}\right\}$ with $k = 1$ referring to the lowest ranked and $k = 4$ referring to the highest ranked. Retain the estimate$\hat{\beta}_{(k)}$ with the largest index k subject to $0 < \hat{\gamma}_{(k)} < 1$.

This estimation procedure works because of the large geographical coverage of the Burbio and R2L data, and the large set of schools for which the Safegraph-NCES visits data is neither too noisy or sparse. Not only is $0 < \hat{\gamma}_{(k)} < 1$ always satisfied for at least one set of estimates, but in many instances we find that $0 < \hat{\gamma}_{(4)} < 1$, which enables us to use the estimates with the highest R-squared. The retained regressions generally provide a close fit, with an interquartile range of R-squareds of 0.68 to 0.96.

There are two states, however, for which the estimation is less successful. In Arkansas, the schooling modes from R2L and Burbio data vary little over time, which implies R-squared close to zero. Similarly, in Maine, the best fitting R-squared is 0.23, except for one CBSA. For those, we assume that the relationship between EIPL and school visits is the same as in neighboring states and compute EIPL with estimates of $\hat{\beta}$ from states of the West South Central division for Arkansas and states of the New England division for Maine. As support for this assumption, note that there are three multi-state CBSAs with schools in Arkansas, respectively Maine, that have larger variations in schooling mode and, as a result, regression R-squared higher than 0.25.[4] The retained $\hat{\beta}$ estimates for these three CBSAs turn out to be similar to the ones from neighboring states that we use to compute EIPL for the rest of Arkansas and Maine.

[2] Burbio and R2L data offer a better geographic coverage and data at the weekly frequency. When the sample period used to estimate Equation (A.3) includes the weeks of August and early September, the panel based on R2L data is unbalanced in some regions because R2L's earliest tracking date differs across school districts. The panel is always balanced when working with the Burbio data.

[3] In practice, we find that the best fitting sample periods are mostly during Fall 2020. This is because in Spring 2021, most schools had stopped fully virtual schooling and thus $T_{c,t} \approx 100 - H_{c,t}$. Since $\Delta\widetilde{v}_{c,t}$ is subject to idiosyncratic variation, this then implies that $\hat{\gamma}_c \to 1$ and $\hat{\beta}_c \to 0$ with a R-squared that tends to zero.

[4] Fayetteville-Springdale-Rogers and Memphis in Arkansas, and Portland-South Portland-Biddeford in Maine.

B Data description

B.1 Schooling mode trackers

To identify prominent schooling mode trackers, we search Scopus, Mendeley, Center for Open Science, and the ICPSR data repository using the keywords "COVID-19", "School closure", "School reopening", and "Schooling mode". Wethen manually check the first 100 papers matching those keywords on Google scholar. In addition, we search on Google as well as on several education blogs to identify trackers that may not have been used in academic research. We end up with eight trackers:

- Burbio is a private company specialized in aggregating school, government, library and community event information. Burbio publishes a weekly School Opening Tracker for almost all U.S. counties based on information from 1,200 public school districts representing 47% of U.S. public K-12 student enrollment in over 35,000 schools.[5]

- The Center on Reinventing Public Education (CRPE) is a nonpartisan research and policy analysis center affiliated with the Arizona State University. The CRPE data is a product of the Center's "Evidence project" that contains data for 477 school districts representing about 20% of U.S. public K-12 student enrollment. School district weights were designed by the RAND corporation to make the CRPE sample representative.

- The COVID-19 School Data Hub (CSDH) was assembled by submitting data requests to state education agencies for their record of learning models used by schools and districts during the 2020-21 school year. The school-level data, which contains data for almost 60,000 schools, covers about 2/3 of U.S. public K-12 student enrollment. The district-level data, which contains data for 5,000 school districts, covers about 30 percent of U.S. public K-12 student enrollment. The school-level and district-level data partly overlap with each other.

- Education Week is an independent news organization owned by a nonprofit educational organization called Editorial Projects in Education. The EdWeek tracker is a compilation of state-level orders or recommendations and public statements or actions from governors and state officials.

- The Elementary School Operating Status database (ESOS) provides data on school districts' primary operating status in the first and last grading period of the 2020-21 school year. Data are available for 9,195 *elementary* school districts, and is a near universe of all elementary students who account for 45% of total U.S. public K-12 student enrollment.

- The Institute of Education Sciences is the independent, non-partisan statistics, research, and evaluation arm of the U.S. Department of Education. Its monthly School Survey Dashboard, started in January 2021 and ending in May 2021, provides data at the state level in 46 states, by collecting data from between 4,000 and 4,500 schools.[6] The sample is stratified to ensure that it is representative across regions of the country and type of location of the schools.

- MCH strategic data is a private company that compiles institutional and marketing data, and specializes in data collection for the segments of Education and Health Care. The MCH tracker is a near-universe of school district's operational status, which includes not only the teaching methods in place in each school district but also the student and staff mask policies and the availability of COVID testing on- and off-site. Data is continuously collected through the year and does not have a clear time frame (other than referring to a specific Semester of the school year).

[5] These figures refer to Burbio's methodology for the 2020-21 school year. Burbio's sample size increased to 5,000 school districts (covering 70% of U.S. K-12 student enrollment) for the school opening tracker of the 2021-22 school year.

[6] About 6,000 schools are sampled, but not all the schools responded in the survey. The numbers are lower for the first round of the survey in January 2021: 3,300 out of the 5,000 schools responded, providing data for 42 states.

- Return2Learn (R2L) is a schooling mode tracker constructed by the American Enterprise Institute and Davidson College. The R2L data consists of weekly indicators from August 2020 onward of the share of public school students engaged in one of the three learning modes. The data is available at the school district level, covering about 8,000 districts in over 3,000 counties that account for about 90% of U.S. public K-12 student enrollment.

Table B1 overviews the eight trackers. There are a number of key differences between them. First, the time frequency, geographical coverage, and level of aggregation vary. On the one hand, CSDH offers the highest level of disaggregation, in the sense that it includes data at the school level, but only a subset of the CSDH schools have information available at the weekly frequency (about 10,000 schools). R2L, Burbio and EdWeek, on the other hand, provide data at the weekly frequency, but only at the district-, county-, and state-level, respectively. Second, the data collection methods are different, and as a result the degree of systematicity is not uniform across trackers. Some collected data at a lower cost (e.g. web scrapping) to increase coverage and hence representativeness; others, such as the CRPE, selected a smaller set of school districts and calculated sampling weights to extrapolate statistics from the selected districts. Third, with the exception of EdWeek, the trackers agree on the choice of the measured items – whether a school offers mostly in-person, hybrid, or remote learning – but there are important differences in how each tracker defines these indicators (see the next section).

Table B1: Overview of schooling mode trackers

Data structure and coverage	Data source and collection method	Measures
Burbio		
• Balanced panel • Weekly data spanning the 2020-21 school year • 3,214 counties (aggregation of data collected from 1,200 school districts)	• Web scraping of school district websites, local news reports, social media, and other publicly available information. • Use the most in-person option available to the general student population to assign a learning mode to the school district.	% of school districts (weighted by student enrollment) within a county that operate in either In-Person, Hybrid, or Remote learning
CRPE (Center on Reinventing Public Education)		
• Panel data sampled at irregular time intervals • Three point-in-time data collection during the Summer and Fall term of 2020 (Jul.26 - Aug.1, Aug.16 - Aug.22, Nov.1 - Nov.7) • 477 school districts	• Web scraping of school district websites, local news reports, social media, and other publicly available information • CRPE data comes with assigned district weights created by the RAND corporation to create a representative sample	0/1 indicators of either In-Person, Hybrid, or Remote learning
COVID-19 School Data Hub (CSDH)		
Mixed levels and data frequencies: • Weekly: 10,121 sch. / 3,301 dist. • Bi-weekly: 4,725 sch. / 540 dist. • Monthly: 33,086 sch. / 1,380 dist. • Quarterly: 144 districts • Bi-annual: 11,928 sch.	• Data requests submitted to state education agencies for their record of learning models used during the 2020-21 school year. Data requested at either the school or district level, as available by the state, and at the most frequent reporting intervals available.	0/1 indicators of either In-Person, Hybrid, or Remote learning

Education Week (EdWeek)		
• Balanced panel • Weekly data spanning the 2020-21 school year • 50 states, the District of Columbia and Puerto Rico	• Information gathered from orders or recommendations issued at the state level, and public statements or actions from governors and state officials. State order may be subject to waivers or overridden by other officials.	0/1 indicators for multiple categories: Full closure (and whether in effect or not), Partial closure, Ordered open, No order in effect, Some grades ordered open, Only hybrid or remote instruction
Elementary School Operating Status database (ESOS)		
• Panel data sampled at irregular time intervals • Two point-in-time data collection: Sep.20-30, 2020 and Apr.20-30, 2021 • 9,195 elementary school districts	• Information gathered from elementary school reopening plans broadly available to the public as parents and local communities.	0/1 indicators of either In-Person, Hybrid, or Remote learning, with several options for Hybrid learning (part day / part week / rotating weeks / other)
School Survey Dashboard of the Institute of Education Sciences (IES-SSD)		
• Unbalanced panel (some states or jurisdictions do not participate or do not meet the minimum participation guidelines for reporting in all waves) • Monthly frequency from January through May 2021 • 50 states, the District of Columbia and Puerto Rico	• Survey administered through a web-based data collection system in jurisdictions that have agreed to participate. Intended survey respondents are school or district test coordinators (State coordinators also invited to respond to individual school surveys or submit results for many schools at once).	% of student enrolled in either In-Person, Hybrid, or Remote learning
MCH strategic data		
Two cross-sectional datasets for the 2020-21 school year: • Fall 2020: 14,893 school districts • Spring 2021: 16,727 school districts	• Proprietary data compilation process and scoring method, with is continuous data updated throughout the school year.	0/1 indicators of either In-Person, Hybrid, or Remote learning, with several options for In-Person (full / on premises) and Hybrid learning (full / partial)
Return2Learn (R2L)		
• Balanced panel • Weekly data spanning the 2020-21 school year • 8,608 school districts	• Web scraping of school district websites, local news reports, social media, and other publicly available information. • Weekly updates of the data using a machine learning approach to analyze whether the new content indicates a change in operational status.	0/1 indicators of either In-Person, Hybrid, or Remote learning

Notes: The table describes the structure, coverage, data source and collection methods, and measurements provided in eight publicly-available learning mode trackers.

Figure 1 in the paper shows the student-weighted average of each learning mode from each tracker. For CSDH, we average data using school-level and district-level student enrollment data. For schools districts that also have school-level data available in CSDH, we first aggregate data to the district level by taking the (student-weighted) average of the district-level and school-level indicators. For EdWeek, data for hybrid and remote learning is discontinued over the sample period (panels (b) and (c) of Figure 1). For MCH, the data is collected at various, potentially irregular, point in time within each semester (making it difficult to assess whether MCH data match the timing of changes across learning modes); the MCH data in Figure 1 are plotted at the midpoint of the Autumn and Spring semesters.

Definition of In-Person / Hybrid / Remote learning

Table B2 complements Table B1 by reporting the definition of In-Person, Hybrid, and Remote learning implemented in the eight pandemic schooling mode trackers covered by our analysis.

Table B2: Definitions of main concepts (In-Person / Hybrid / Remote) by schooling mode tracker

In-Person	Hybrid	Remote
Burbio		
Students attend in-person every day.	Students are divided into cohorts and attend 2-3 days in-person and 2-3 days virtually.	Residual category.
CRPE (Center on Reinventing Public Education)		
Schools open with only in-person instruction (no virtual/remote instruction) for at least one grade band.	Schools open with some combination of in-person and virtual/remote instruction for at least one grade band.	Schools open with only virtual/remote instruction (no in-person instruction) for at least one grade band.
COVID-19 School Data Hub (CSDH)		
Fully in-person instruction 5 days a week for all or most students.	A blend or combination of in-person and virtual instruction for all or the majority of students.	Fully remote or distance learning for all or the majority of students.
Education Week (EdWeek)		
In-person instruction must be available to all students, or available for certain grade levels, either full- or part-time.	Full-time in-person instruction is either not allowed in certain regions of the state or is only available for certain age groups. Hybrid instruction may be allowed.	In-person instruction is not allowed.
Elementary School Operating Status database (ESOS)		
Students attend in person at least 4 full days per week. Schedules may be shorter than traditional hours but longer than part time (at least 4 hours).	• Hybrid (part day) : In-person learning with part-time or significantly reduced hours per day (4 hours or fewer). • Hybrid (part week) : In-person learning 1, 2 or 3 days per week. • Hybrid (rotating weeks) : In-person learning in alternating weeks. • Hybrid (other): Any other hybrid plan not previously specified or a combination of multiple hybrid plans (e.g., part day and part week).	Students are not allowed to attend school in person. • School districts are counted as fully remote even if they allow in-person attendance with limited exceptions (e.g., English Learners, students with disabilities). • If a minority of grades are allowed in-person attendance (example: K-1 in person and grades 2-6 remote), districts are counted as remote based on a majority of grades being remote.

School Survey Dashboard of the Institute of Education Sciences (IES-SSD)		
Open with full-time in-person instruction.	Open with both remote/online and in-person instruction. If chosen, another question asks for type of hybrid (part of the week, every other week, every three weeks, other).	Remote or online instruction only.
MCH strategic data		
School district offers face-to-face instruction 5 days per week to all students at all available grade levels.	School district offers face-to-face instruction but less than 5 days a week, or to a subset of students.	School district offers no face-to-face instruction and learning is conducted online to all students at all available grade levels.
Return2Learn (R2L)		
All grade levels can attend school in buildings 5 days per week, though families can opt for fully remote instruction or a hybrid model.	Either students in some grades can return to buildings in person while other grades can only return in a hybrid or remote model or all students can return to buildings for 4 days or less each week (or 5 partial days) while learning remotely from home the remaining time.	All grade levels above first grade participate in virtual instruction 5 days per week, with no option for in-person or hybrid learning. Districts that only allowed in-person or hybrid instruction for prekindergarten, kindergarten, first grade, or select subgroups of students are included in this category.

Notes: The table describes the structure, coverage, data source and collection methods, and measurements provided in eight publicly-available schooling mode trackers.

There are several interesting common patterns regarding the definitions and measurement of in-person / hybrid / remote learning. First, most school trackers opt for a rule that favors the most in-person available option (which may help reduce differences in measured learning modes across the different trackers). For instance, according to Burbio's documentation, if a district offers both traditional and virtual options, the district is categorized as "Traditional". Similarly, in ESOS the rules are that "Districts are counted as in person even if some students opt out due to parental preference or high-risk medical conditions. If multiple options are provided and parents can select an option (e.g., in person, hybrid, remote), schools are coded by the most generous in-person option provided.". Second, several school trackers include some additional categories to describe the type of learning mode offered to the student population. For example, CRPE includes the descriptors: "Varies by school", "Varies by grade band", "To be announced", "No information". The use of additional categories to complement in-person / hybrid / remote learning illustrate well the challenges of describing school modalities during the pandemic. Consider for instance the ESOS definitions: They show not only the variety of the hybrid learning modes that were in place (part day, part week, rotating weeks, etc.), but also the possibility of a number of important exceptions (in-person attendance for certain groups of students, different rules across grades, etc.) that make it difficult to assign a single learning mode to a school. Note that this is potentially a source of additional noise through frequent changes in the operational status of schools.

B.2 NCES data

The U.S. Department of Education's NCES is the primary federal entity for collecting and analyzing data related to education. The NCES regularly publishes statistics on both public and private schools and also makes available different datasets on individual schools. We mainly make use of two NCES datasets:

- Common Core of Data (CCD; see https://nces.ed.gov/ccd/). CCD is a comprehensive annual database of all public elementary and secondary schools and school districts (including public charter schools). The CCD consists of five surveys completed annually by state education departments from their administrative records. The information includes a general description of schools and school districts, including name, address, and phone number; number of students and staff, demographics (including the gender and racial makeup of the schools students); and fiscal data, including revenues and current expenditures. We use the 2020-2021 CCD school data files released in January 2022.

- The second dataset is the Private School Universe Survey (PSS; see https://nces.ed.gov/surveys/pss/), which is a biennial survey that collects data on private schools and serves as a sampling frame for other NCES surveys of private schools. The PSS data include a general description of schools, teachers, and students (including the gender and racial makeup of the schools students) in the survey universe. The schools surveyed in the PSS come with a survey weight. We use the 2019-2020 data files released in February 2022.

Table B3 compares aggregates from the CCD and PSS to the NCES's digest of education's statistics (see https://nces.ed.gov/programs/digest/). The CCD files we are using were released only recently and have not yet been used by the NCES to produce official statistics, but the close similarity between all counts (number of educational institutions, number of students, number of teachers) suggests that the CCD and PSS files put together cover the universe of elementary and secondary schools.

We complement the NCES datasets with information from the Education Demographic and Geographic Estimates (EDGE; see https://nces.ed.gov/programs/edge/). EDGE is a program run by the NCES to create and assign address geocodes (estimated latitude/longitude values) and other geographic indicators to public schools, public local education agencies, private schools, and post-secondary schools, and create area-type indicators (City, Suburban, Town, and Rural). We use the 2020-2021 geocodes to improve the reliability of the match between the CCD/PSS files and Safegraph data, and the area-type indicators to assess the relation between EIPL and local school characteristics (see Figure C2).

B.3 Safegraph

Safegraph is a data company that aggregates anonymized location data from cell phone applications in order to provide insights about foot traffic (visits) to physical places, called Places of Interest (POI). Each POI in Safegraph's data is identified by a unique persistent `placekey` identifier. Details of the spatial hierarchy of POIs are important to understand visit attribution. A POI is a polygon, and some of the polygons are encompassed into larger polygons. When it so happens, the "child" polygon receives a `parent_placekey` equal to the `placekey` of the encompassing "parent" POI.

About 99% of POIs come with a 6-digit industry NAICS code, and about 80-85% of Safegraph's POIs come with information on visits.[7] In our analysis of POIs with NAICS 611110 ("Elementary and Secondary Schools"), about 5% have a `parent_ placekey`, which is almost always shared with a POI that is classified as NAICS 624410 ("Child and Youth Services") or NAICS 813110 ("Religious organizations"). To reduce noise in the visits data, we aggregate up these visits and attribute them to the school that is paired to these non-611110 NAICS POIs.

As described in the main text, we match POIs by school name and geolocation or address to the universe of public and private schools from the NCES' CCD and PSS files. The details of the matching procedure as well as other details of our Safegraph visits data construction are explained in a supplementary file posted on the authors' websites.

[7]See https://docs.safegraph.com/docs/core-places#section-naics-code-top-category-sub-category for information on Safegraph's algorithm for attributing NAICS codes to the POIs covered by the Core places dataset.

Table B3: Comparison with the NCES digest of education's statistics

Number of educational institutions		
	NCES table 105.50 (1)	CCD & PSS (2)
Public Schools	**98,469**	**101,688**
Elementary	67,408	68,953
Secondary	23,882	21,434
Combined	6,278	6,678
Other[a]	901	4,623
Private Schools	**32,461**	**27,641**
Elementary	20,090	17,378
Secondary	2,845	2,301
Combined	9,526	7,962
All	**130,930**	**129,329**
Number of students (in 1,000s)		
	NCES table 105.20 (1)	CCD & PSS (2)
Public Schools[b]	**50,686**	**50,834**
Prekindergarten to grade 8	35,496	33,415
Grades 9 to 12	15,190	17,419
Private Schools	**5,720**	**4,090**
Prekindergarten to grade 8	4,252	3,450
Grades 9 to 12	1,468	0.639
All	**56,406**	**54,924**
Number of teachers (in 1,000s, full-time equivalents)		
	NCES table 105.40 (1)	CCD & PSS (2)
Public Schools	3,170	2,911
Private Schools	482	401
All	**3,652**	**3,312**

Notes: NCES numbers refer to the year 2017-2018 (most recent release of the NCES's tables covering both public and private schools for the same school year). Public schools classified as "Other", denoted by [a], includes special education, alternative, and other public schools not classified by grade span. NCES enrollment numbers in public schools, denoted by [b], include imputations for public school prekindergarten enrollment in California and Oregon.

Table B4 describes the outcomes of our match algorithm. We obtain direct merges for about 75,000 schools, which represent more than 60% of schools in the NCES files. Matching based on names and GPS coordinates yields an additional 35,000 matches of high quality, bringing the total of matched schools to 110,644 schools (93,312 public schools and 17,332 private schools). We exclude the remaining schools because either the match is not by school name, the match is of poor quality, or there is no match at all.

Out of the 110,644 matched schools, we discard about 37,000 due to sparse or noisy visits data (see the supplementary file for details). The remaining 73,194 schools constitute the "in-scope" dataset that we use for the estimation of EIPL. Table B5 compares different observables in the full CCD-PSS file (columns 1 and 4) with the schools that we match to Safegraph (columns 2 and 5), and the in-scope dataset (columns 3 and 6). The characteristics of the universe of schools in the CCD-PSS file and the matched dataset are very similar. The characteristics of the in–scope dataset, by contrast, are somewhat different. We therefore construct weights for each school to make the in-scope dataset representative (see the supplementary file for details). As the table shows, once we apply these weights, the in-scope dataset is representative of the universe of schools in the CCD-PSS file.

Table B4: Results of matching Safegraph with NCES schools

	Number of schools (1)	% of the NCES schools (2)
Merge on name/address/zip-code	62,701	50.9
Merge on name/address	12,411	10.1
Match on name/lat/lon, high quality[a]	34,116	27.7
Fuzzy match on name/address/zip-code, high quality[b]	579	0.47
Fuzzy match on name/zip-code, high quality[b]	837	0.68
Fuzzy match on address/zip-code, high quality[b]	2,585	2.10
Match on name/lat/lon, low quality[a]	8,939	7.25
Fuzzy match (any combination), low quality[b]	27	0.02
Not matched	1,039	0.84
Total	123,234	100

Notes: The table reports the counts and share of schools from the NCES's CCD and PSS files that we merge or match to Safegraph at the different steps of the algorithm. High quality as denoted by [a] refers to schools that are closest to each other within the area defined by GPS coordinates rounded to the first decimal place, and the additional requirements that they are less than 250 meters apart and the Levenshtein distance between school names is under 0.250. High quality as denoted by [b] refers to schools that receive a matching score higher than 0.85 through Stata's fuzzy name matching command.

Table B5: Comparison between all schools and schools from the in-scope dataset

	Public schools			Private schools		
	All (1)	Matched[a] (2)	In scope[b] (3)	All (4)	Matched[a] (5)	In scope[b] (6)
Sample count	101,662	93,312	63,395	21,572	17,332	9,832
Student-teacher ratio	16.0	15.6	15.5	10.3	10.2	10.7
% Male	52.1	52.1	51.8	52.5	52.0	51.8
% Indian	1.84	1.74	1.37	0.70	0.70	0.61
% Asian	3.95	3.96	4.13	6.15	5.95	6.20
% Pacific	0.40	0.34	0.35	0.58	0.61	0.62
% Hispanic	25.6	25.3	24.6	12.0	11.7	13.1
% White	49.3	50.1	51.8	64.4	65.9	64.0
% Black	14.5	14.1	13.3	11.6	10.4	10.5
% Other	4.44	4.46	4.40	4.64	4.66	4.99
% Free lunch[c]	39.2	39.1	38.7	n.a.	n.a.	n.a.
% Reduced-price lunch[c]	3.70	3.72	3.85	n.a.	n.a.	n.a.
City	27.5	26.5	26.9	34.3	34.4	39.6
Suburban	31.4	31.7	28.1	38.9	37.4	38.5
Town	13.2	13.5	14.6	8.06	9.12	8.91
Rural	27.9	28.3	30.4	18.7	19.1	13.0

Notes: Schools marked as "Matched", denoted by [a], refer to schools matched to Safegraph data. Schools marked as "In scope", denoted by [b], refer to schools matched to Safegraph with visits data neither too sparse or too noisy. Except for the sample count, all the statistics for the "In scope" data are computed using school weights. % Free lunch and % Reduced-price lunch, denoted by [c], refer to the school shares of students who are eligible for free and reduced-price lunches, respectively.

B.4 Data used for the regressions

ACS data The socio-demographic and income variables are based on the American Community Survey (ACS) 5-year estimates for the release years 2016-2019. The estimates are computed at the Census Block Group (CBG) level. To aggregate data to the 5-digit zip-code level, we use the ZIP-TRACT crosswalk provided by the U.S. Housing and Urban Development (HUD)'s Office of Policy Development and Research (see https://www.huduser.gov/portal/datasets/usps_crosswalk.html). To measure population density, we use land area data from the U.S. Department of Agriculture (USDA; see https://www.ers.usda.gov/data-products/atlas-of-rural-and-small-town-america/). We also use the USDA's Rural-Urban continuum codes.[8]

EDGE data Data on the local labor costs of hiring PK-12 educators come from the Education Demographic and Geographic Estimates (EDGE; see https://nces.ed.gov/programs/edge/). The index proxies the outside options of PK-12 educators by using local information (obtained from restricted-use data from the ACS) on the wages and salaries of comparable workers, while excluding from the estimation sample anyone who has a teaching or educational administration occupation or who is employed in the elementary and secondary education industry.[9] We also use the school neighborhood poverty estimates from EDGE in robustness checks presented in the supplementary file.

CPS data Data on teachers' unionization rates are computed from the Current Population Survey (CPS; see https://www.census.gov/programs-surveys/cps.html) using the outgoing rotation group samples of the survey. We pool together data from 2018 through 2020 for teachers and instructors in elementary and secondary schools (PEIO1OCD 2310, 2320, 2330, 2340) and define the unionization rate as the share of teachers who are either members of a labor union or covered by a union. We aggregate unionization rates to the CBSA level or to the state level for missing CBSAs. We validate the state-level unionization rates against official tabulations of state-level unionization rates of public school teachers published by the NCES (see https://nces.ed.gov/surveys/sass/tables/sass0708_043_t1s.asp).

SEDA data Data on school-level and district-level test scores come from the Stanford Education Data Archive (SEDA; see http://purl.stanford.edu/db586ns4974), version 4.1. SEDA is a data initiative whose goal is to provide nationally comparable, publicly available test score data for U.S. public schools and public school districts. In our main analysis, we use test scores at the levels of school districts due to wider data coverage. These test scores are pooled across all school grades and across years 2013 through 2018; they are normalized either through a cutscore standardized to the nationally averaged reference cohort within subject, grade, and year (the CS scale), or through a grade-cohort standardized score (GCS scale); they are available separately for mathematics and RLA. We use the CS scale and take the mean of the mathematics and RLA test scores. At the school level, test scores are not available by subject and they cover a smaller portion of our dataset.

NERD$ data Data on spending per student at the school level come from NERD$, the National Education Resource Database on Schools. NERD$ is a data initiative of the Edunomics lab and the Massive Data Institute at Georgetown University (see https://edunomicslab.org/) that gathers together school spending data that tend to be scattered across different states' websites. The data we use are from the update of NERD$ dated from January 28th, 2022 and contain school-by-school actual spending amounts for the year of 2018-2019. The data matches 94% of the public schools of our dataset.

ESSER data Data on ESSER funding come from the compilation put together by Return2Learn and available on R2L's website (see https://www.returntolearntracker.net/esser/). The raw data covers all three waves of ESSER. Data are available at the level of school districts, and the R2L database comes

[8]See https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx.

[9]"Comparable" means that the index controls for a host of socio-demographic and employment characteristics on individuals who are college graduates; see https://nces.ed.gov/programs/edge/Economic/TeacherWage.

with the NCES identifier for school districts; when matched to our own, it covers about 91% of school districts that include 95% of the public schools in our dataset.

COVID data Data for COVID cases and deaths at the county level are based on the daily count and rates from the New York Times, the Johns Hopkins Coronavirus Resource Center, and the Centers for Disease Control and Prevention (CDC). County-level COVID vaccinations at the are daily rates are from the CDC. To aggregate to the weekly level, we take the mean of the daily values for each variable. County-level counts of ICU beds come a report from Kaiser Health News accessed through a compilation available at: https://github.com/JieYingWu/COVID-19_US_County-level_Summaries/tree/master/data.

Election data County-level results for presidential elections are downloaded from the MIT election Data and Science Lab (see https://electionlab.mit.edu/data). We use results for the 2020 presidential elections in our main analysis and results for the 2016 presidential elections in robustness analyses.

NPIs data Data at the county-week level on Non-Pharmaceutical Interventions (NPIs) come from the repository of the Centers for Disease Control and Prevention (CDC; see https://data.cdc.gov/). We use information about the following NPIs: 1) Stay-at-home orders, which can be advisory/recommendation, mandatory only for individuals in certain areas of the jurisdiction, mandatory for at-risk individuals, or mandatory for all individuals, 2) Gathering bans, which can be bans on gatherings of more than 100 persons, more than 50 persons, more than 25 persons, more than 10 persons, or all social/public gatherings, 3) Mask mandates, which is an indicator that takes the value of 1 when a mask is required in public and is 0 otherwise. All series are from the September 10, 2021 update of the CDC data.

OA data In robustness checks presented in the supplementary file, we use data from the Opportunity Atlas (OA; see https://www.opportunityatlas.org/). OA provides access to social mobility data assembled by researchers from the Census Bureau, Harvard University, and Brown University. The data we use is at the level of Census tracts, which we can link to our dataset through the Census Block Group (obtained through Safegraph) of each school. We use data pooled across genders and races/ethnicities.

Descriptive statistics Table B6 presents descriptive statistics of the main variables of interest, weighted by the student population of the public and private schools included in our analysis.

To complement the multivariate regression results in the main text, Figure B1 presents pairwise correlations coefficients between the different variables. Household income, share of adults with college or higher education, and share of dual-headed households are strongly correlated with each other. In contrast, a school's share of non-white students is essentially uncorrelated with income and education in the same zip code. Also note that district-level ESSER funding per student is negatively correlated with income, education, share of dual-headed households, and school test scores (strongly so), but negatively correlated with the school share of non-white students.

C Detailed results for EIPL

Here we provide details for the analysis of EIPL in Section 4 of the main text. Additional results are provided in a supplementary file posted on the authors' websites.

C.1 EIPL by school type and grade

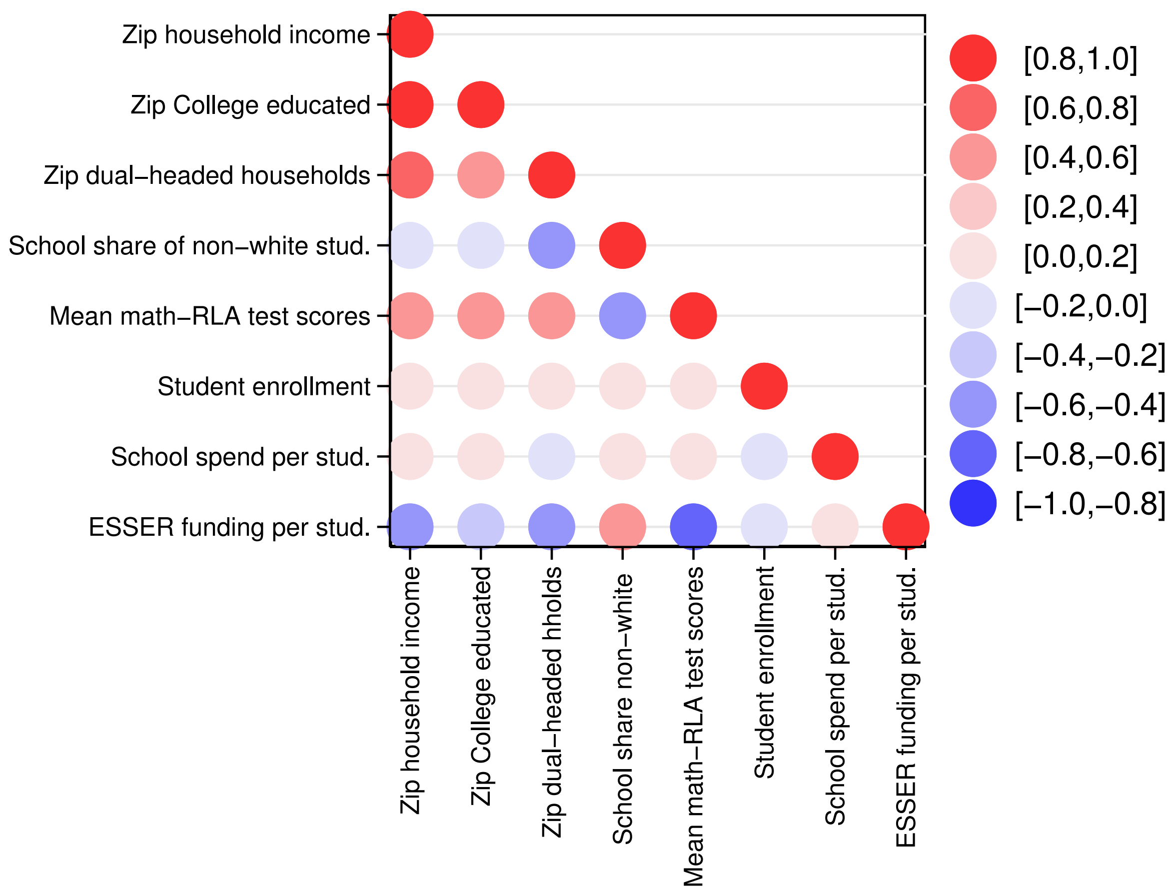
Panel (a) of Figure C1 shows differences in average EIPL by school type and time period. During the first three months of the pandemic, there is almost no difference in EIPL across school types. During both Fall 2020 and Winter/Spring 2021, however, we see substantial differences. Over the entire 2020-21 school year, EIPL is 10% lower for public schools than for private schools, with public charter schools averaging the least EIPL, followed by public non-charter, private non-religious, and private religious schools.

Table B6: Descriptive statistics of the school-level regression variables

	Mean	St. Dev.	Percentile 25th	Percentile 50th	Percentile 75th	Min.	Max.
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
				(a) Public schools			
Zip-level household income	75,673	30,459	55,904	67,712	87,191	7,770	432,067
Zip-level share of college educated	0.27	0.15	0.16	0.23	0.36	0.01	0.93
Zip-level share of dual-headed households	0.70	0.14	0.62	0.72	0.80	0.00	1.00
School share of non-white students	0.48	0.32	0.19	0.43	0.77	0.00	1.00
Mean math-RLA test scores	0.01	0.32	-0.20	0.01	0.21	-1.25	1.25
Student enrollment	576	460	317	467	682	6	21,049
School spending per student	12,463	4,488	9,526	11,502	14,285	171	49,957
ESSER funding per student	3,238	2,470	1,498	2,795	4,197	0	30,189
				(b) Private schools			
Zip-level household income	86,915	39,598	60,143	76,082	103,059	22,512	397,509
Zip-level share of college educated	0.36	0.18	0.20	0.31	0.48	0.02	0.91
Zip-level share of dual-headed households	0.70	0.15	0.62	0.72	0.81	0.05	0.99
School share of non-white students	0.36	0.30	0.12	0.27	0.53	0.00	1.00
Student enrollment	228	252	65	155	297	6	3,825

Notes: The table reports the mean, standard deviation (St. Dev.), the 25th, 50th, 75th percentiles, and the minimum (Min.) and maximum (Max) values of the right-hand side variables of the school-level regressions. All statistics are computed using school weights.

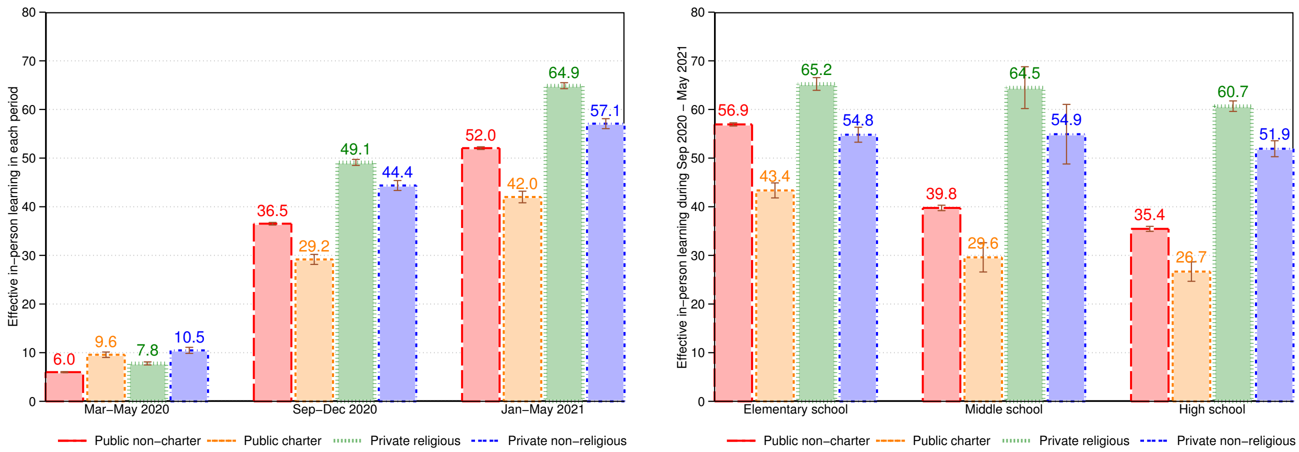
Figure B1: Cross-correlations of the school-level regression variables



Notes: The figure show the cross-correlations of the variables used in the school-level regressions. Correlations are computed using school weights.

Figure C1: Effective in-person learning by school type and grade

(a) School type by time period **(b)** School type by school grade



Notes: The figures show student-weighted average EIPL for private schools versus public schools by time period and by school grade.

Panel (b) of Figure C1 reports on differences in average EIPL between September 2020 and May 2021 by school type and school grade. Across all four school types, EIPL is highest for elementary schools and lowest for high schools. For private schools, the difference in EIPL across school grades is smaller than for public schools.[10] In other words, the differences in EIPL between public and and private schools that we observe in Panel (a) are in large part due to differences in EIPL at the middle and high school level.

As shown in Figure C2, the ranking of EIPL by school type and grade holds true within all four regions of the U.S. Interestingly, the relation is weaker in the South for private schools; and the difference between public non-charter and public charter schools is reversed in the western part of the country. Figure C2 also shows that the magnitude of the EIPL gap between public and private schools differs across regions; for instance it is larger in cities of the Northeast region of the country.

C.2 Regression results

Table C1 presents the estimates reported in Figure 4 of the main text.

D Disparities in schooling mode tracker results

D.1 Relation to local, school and regional characteristics

Figures D1a and D1b are the counterparts to Figure 4 in the main text, reporting the results of the holistic regressions for respectively in-person and remote learning. The different markers in Figures D1a and D1b denote the different schooling mode trackers used to run the regressions, namely Burbio, the Center on reinventing public education (CRPE), the COVID-19 school data hub (CSDH), the Elementary school operating status (ESOS) database, MCH strategic data (MCH), and Return2Learn (R2L).[11]

[10] Note that for elementary schools, EIPL is slightly higher for public non-charter schools than for private non-religious schools. This change in ranking of school types compared to the ranking across all school grades is due to geographical differences in the relative prevalence of private non-religious elementary schools.

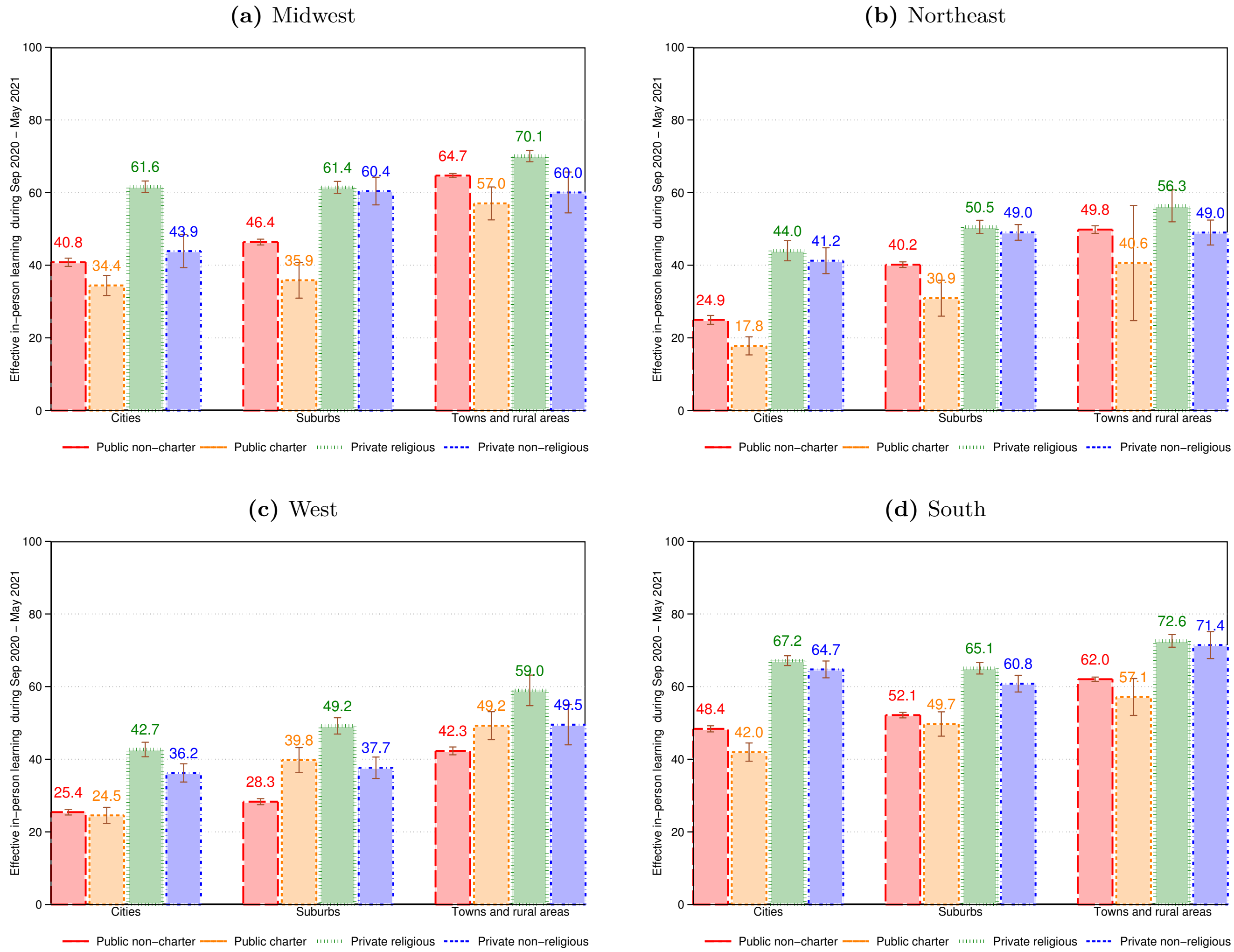
[11] We do not run the regressions on the learning mode trackers that report data only at the state level as this level of aggregation is too coarse to capture the relations evidenced in Figures D1a and D1b. For CSDH data, we run the regressions using the school district data as it is more comparable to the other trackers. Thus, the level of geographic analysis in Figures D1a and D1b is either the school district or the county (for Burbio data). Standard errors are clustered accordingly.

Table C1: Regression results for Figure 4

Dependent variable	Effective in-person learning (EIPL)		
	(1)	**(2)**	**(3)**
Zip-level household income[a]	-4.02***	-6.80***	0.53*
	(0.57)	(0.68)	(0.32)
Zip-level share of College educated[a]	-6.88***	-10.59***	-1.99***
	(0.66)	(0.53)	(0.42)
Zip-level share of dual-headed households[a]	4.05***	0.35	0.44
	(0.61)	(0.74)	(0.40)
School share of non-white students[b]	-20.86***	-15.27***	-6.18***
	(1.20)	(1.34)	(0.86)
Mean maths-RLA test scores[b]	5.10***	5.12***	4.55***
	(0.96)	(0.81)	(0.60)
Student enrollment[b]	-5.21***	-3.56***	-2.92***
	(0.39)	(0.28)	(0.23)
School spending per student[b]	-5.63***	-4.96***	-0.87**
	(0.71)	(0.68)	(0.37)
ESSER funding per student[b]	-5.07***	-0.52	-0.98*
	(0.78)	(0.76)	(0.59)
Teacher unionization rate			-6.90***
			(1.01)
Local index of costs of hiring PK-12 educators			-3.93***
			(1.03)
2020 share of Republican voters			13.45***
			(1.16)
Mask required in public			-7.85***
			(1.08)
COVID vaccination rate (two weeks lag)			8.17***
			(0.56)
School type and grade controls	✔	✔	✔
County health, pop. characteristics, weather, NPIs			✔
R-squared			0.27
# of counties	2,863	2,863	2,633
# of schools	56,632	56,632	56,019
# of school districts	11,391	11,391	10,974

Notes: Each column reports coefficients from a weighted OLS regression on the public school sample, with standard errors clustered at the county level in parentheses and school weights calculated as explained in Appendix E.3. The regressions are estimated on EIPL for the period from September 2020 to May 2021. The school type fixed effects consists of indicators for charter school and non-charter school, and the school grade fixed effects consist of indicators for elementary vs. middle vs. high. vs. combined school for both samples. County health, pop.characteristics, weather, NPIs consist of pre-pandemic ICU bed capacity, two-week lagged county COVID case and death rates, population density in the county, dummies for rural-urban continuum codes, maximum weekly temperature in the county, and dummies for various non-pharmaceutical interventions. The coefficient estimates in column (1) are the results of separate regressions with each one of variable in combination with school type and grade controls. In columns (2) and (3), the coefficient estimates for the affluence measures, denoted by [a], are the result of separate regressions with each one of the measures in combination with the other variables shown in the table. In columns (2) and (3), the coefficient estimates for the school variables, denoted by [b], are the result of regressions where the affluence measures are included jointly (in combination with school type and grade controls in column (2), and in addition with the other regional variables as reported in the table in column (3)).

Figure C2: Effective in-person learning by school type, locale and U.S. region

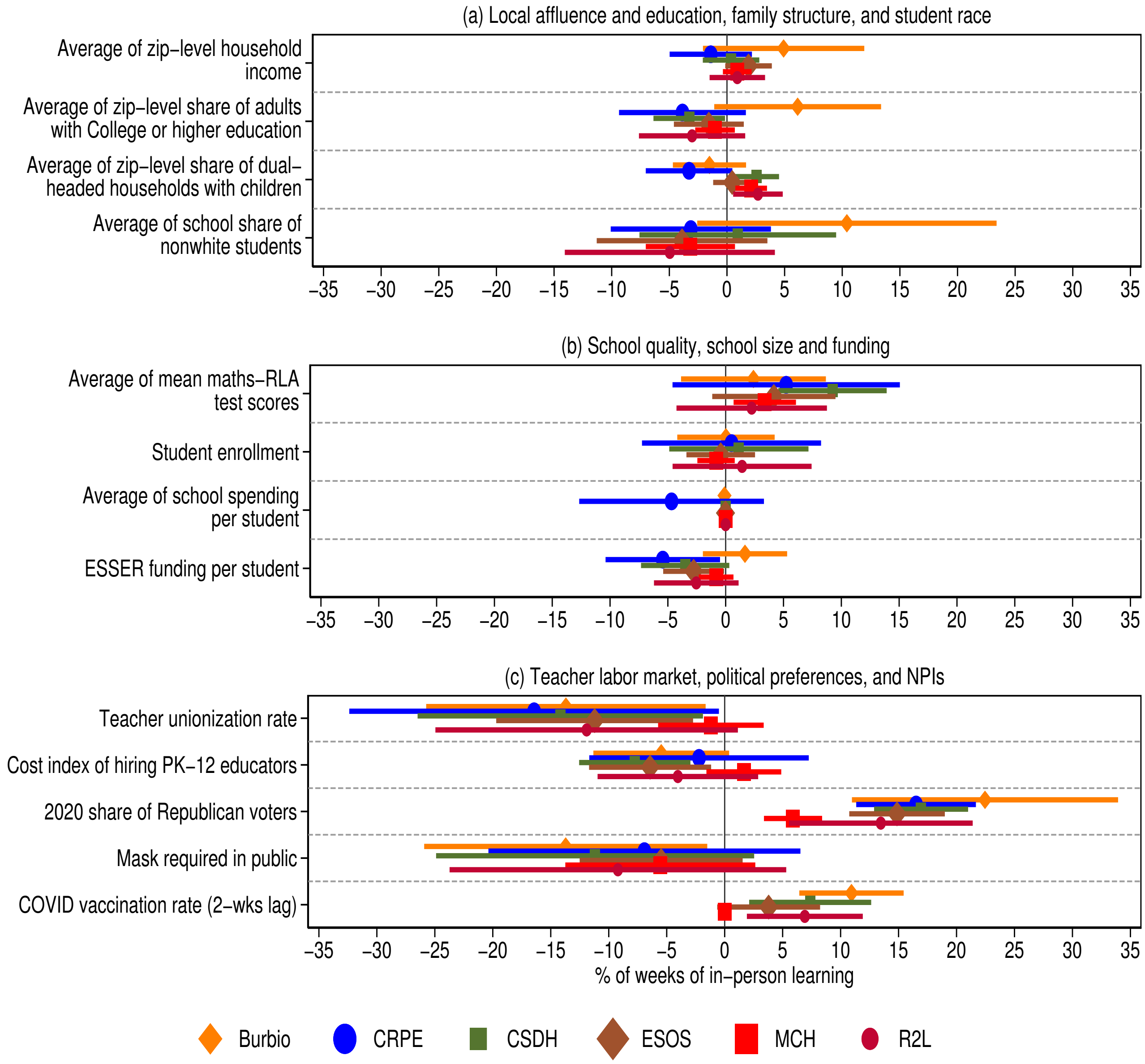


Notes: The figures show student-weighted average EIPL by school type and by locale for the Northeast (CT, MA, ME, NH, RI, VT, NY, NJ, PA), Midwest (IL, IN, MI, OH, WI, IA, KS, MN, MO, NE, ND, SD), South (DE, FL, GA, MD, NC, SC, VA, WV, AL, KY, MS, TN, AR, LA, OK, TX), and West (AZ, CO, ID, MT, NM, NV, UT, WY, CA, OR, WA).

A number of key patterns stand out from Figures D1a and D1b. First, the different school trackers suggest little to no association between in-person or remote learning on the one hand and local affluence, education and family size on the other. This is consistent with the results based on the EIPL dataset after controlling for the full set of school and regional characteristics, but is different from the raw correlations that generally suggest an inverse relation between in-person learning and these local population characteristics.[12] Second, in line with the EIPL dataset, the school trackers show that in-person learning is lower in schools with a larger share of the student body that is non-white. The relationship, however, is not precisely estimated, and in addition the Burbio schooling mode data seems to disagree with the other datasets (albeit with a large standard error). Third, the school trackers point to a positive association between in-person learning and pre-COVID test scores, although the estimates are close to insignificant, and to a negative relation with ESSER funding per student. In the latter case, however, there is some disagreement across the school trackers, especially when we consider the relation between ESSER funding

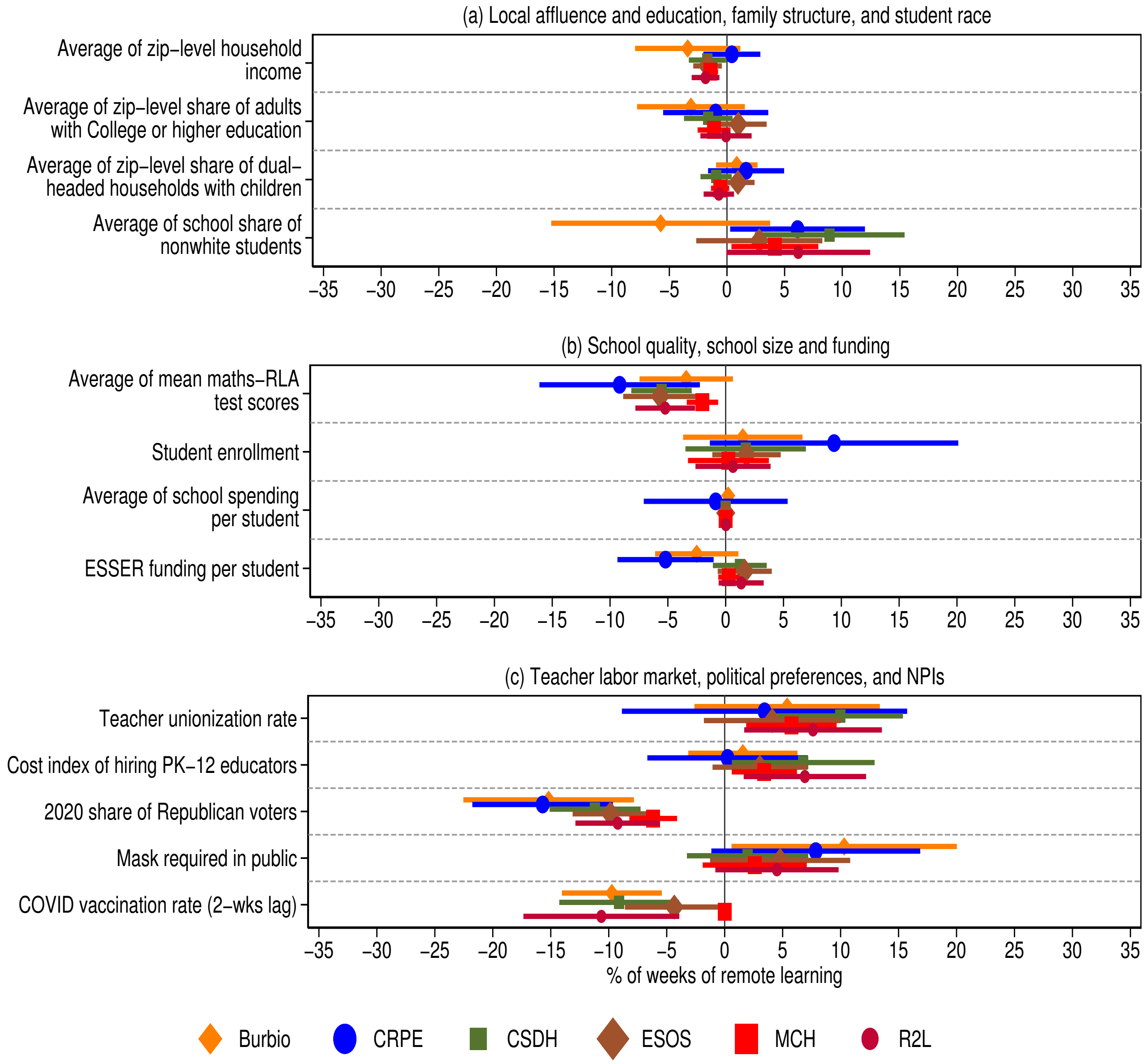
[12] Figures D1a and D1b are also consistent with results based on EIPL in that, if anything, the relation between in-person learning and local characteristics is stronger in what concerns local education compared with local affluence and family size.

Figure D1a: The relationship of in-person learning with school and local characteristics



Notes: The figure shows the estimated effects on in-person learning from OLS regressions based on the following schooling mode trackers: Burbio, the Center on reinventing public education (CRPE), the COVID-19 school data hub (CSDH), the Elementary school operating status (ESOS) database, MCH strategic data (MCH), and Return2Learn (R2L). Standard errors are clustered at either the district (CRPE, CSDH, ESOS, MCH, R2L) or county (Burbio) level. The estimates for the first three variables (local affluence, education, family structure) are the result of separate regressions for each of the three variables in combination with the other variables listed below. The coefficient estimates for the other variables are the result of regressions where all the variables are included jointly. All regressions control for the district or county composition in terms of school type (charter vs. non-charter school) and school grade (elementary vs. middle vs. high. vs. combined school), and control for pre-pandemic ICU bed capacity, two-week lagged county COVID case and death rates, dummies for various other non-pharmaceutical interventions, maximum weekly temperature in the county, county population density, and dummies for rural-urban continuum codes. All estimates except for the "Mask required in public" dummy are scaled to show the implied change in remote learning of going from the 25th to the 75th percentile of the distribution of a variable.

Figure D1b: The relationship of remote learning with school and local characteristics



Notes: The figure shows the estimated effects on remote learning from OLS regressions based on the following schooling mode trackers: Burbio, the Center on reinventing public education (CRPE), the COVID-19 school data hub (CSDH), the Elementary school operating status (ESOS) database, MCH strategic data (MCH), and Return2Learn (R2L). Standard errors are clustered at either the district (CRPE, CSDH, ESOS, MCH, R2L) or county (Burbio) level. The estimates for the first three variables (local affluence, education, family structure) are the result of separate regressions for each of the three variables in combination with the other variables listed below. The coefficient estimates for the other variables are the result of regressions where all the variables are included jointly. All regressions control for the district or county composition in terms of school type (charter vs. non-charter school) and school grade (elementary vs. middle vs. high. vs. combined school), and control for pre-pandemic ICU bed capacity, two-week lagged county COVID case and death rates, dummies for various other non-pharmaceutical interventions, maximum weekly temperature in the county, county population density, and dummies for rural-urban continuum codes. All estimates except for the "Mask required in public" dummy are scaled to show the implied change in remote learning of going from the 25th to the 75th percentile of the distribution of a variable.

and remote learning in Figure D1b, with one third of the point estimates that are below zero. Fourth, turning to the relation with characteristics of the teacher labor market, political preferences, and NPIs, the sign of the point estimates are in general the same as in Figure 4 in the main text. There are two major differences, however: the magnitudes of the effects is very different across datasets, and they lack statistical precision in most instances. Consider for instance the relation between in-person learning and the county-level share of Republican voters in 2020. The magnitude of the relation according to Burbio is four times larger than that based on MCH data. At the same time, the school trackers indicate that in-person learning is negatively related to mask mandates, but the estimates are so imprecise that they suggest effects ranging from 0 to a reduction of in-person learning by 20–25 percent.

To summarize, the regressions presented in Figures D1a and D1b shows results that, on the whole, are consistent with the holistic regression in the main text. However, the levels of statistical association and precision are different, and for some variables the different schooling mode trackers disagree about the sign or magnitude of the impact on in-person learning.

D.2 Regressions with Zearn data

As another concrete example of the usefulness of our EIPL measure, we analyze its relation with students' participation and progress in Math during the pandemic using data from Zearn, an online Math platform used in many public schools across the country. We use Zearn data provided by the Opportunity Insights's economic tracker webpage (see https://tracktherecovery.org/), which provide us with county-level measures of student's participation and progress. We are interested in measuring the statistical association between these two outcomes on the one hand, and in-person learning on the other hand.

Formally, the regression takes the form:

$$y_c = \alpha + \beta_1 T_c + \beta_2 H_c + \delta_{s(c)} + \varepsilon_c, \tag{D.1}$$

where T_c and H_c denote county-level averages of the extent of respectively traditional and hybrid learning that students received during the school year 2020-21; $\delta_{s(c)}$ is a state fixed effect, and y_c is a Zearn outcome (participation or progress in Math) averaged over the period from September 2020 through May 2021. Since Zearn's outcomes are measured in relative terms compared to the pre-pandemic period, and the independent variables T_c and H_c provided by school trackers are in relative terms as well, Equation (D.1) is essentially a diff-in-diff regression with the pandemic as the treatment period and changes in schooling modes as the policy instrument. We seek to compare Equation (D.1) with:

$$y_c = \alpha + \beta EIPL_c + \delta_{s(c)} + \varepsilon_c, \tag{D.2}$$

where $EIPL_c$ is our EIPL measure aggregated to the county-level and averaged over the school year.

Table (D.1) presents the results based on data from Burbio, CSDH, R2L,[13] and our EIPL measure.[14] First, unlike CSDH or R2L, the Burbio data fails to detect any statistically significant relation between learning modes and students' participation or progress in Math. This echoes a key motivation of our analysis, that results obtained from a given tracker may not be robust to the use of an alternative tracker. Second, CSDH and R2L on the one hand and EIPL on the other deliver results that are qualitatively similar, i.e. in-person learning has a positive impact. The quantification based on CSDH and R2L, however, is unclear. Consider for instance the R2L results. If Hybrid consists of 1/3 of in-person learning, then a marginal increase of in-person learning raises participation by $0.0079 + 0.0048 \times 1/3 = 0.0095$ and progress by $0.0069 + 0.0034 \times 1/3 = 0.0080$ standard deviations. These numbers line up with the EIPL estimates, but as the calculation illustrates, they require an assumption (or external estimate) about hybrid learning. Third, the estimates based on CSDH and R2L lack precision. This is not surprising, given

[13] We limit our analysis to these three trackers as they provide a better coverage of the school year 2020-21.

[14] Notice that the number of counties used in these regressions is typically lower than that in the regressions of the main text. This difference is coming from the coverage of Zearn which is more narrow than that of our EIPL dataset.

Table D1: Relationship between Zearn outcomes and in-person learning

Dependent variable	**(a)** Online participation				**(b)** Progress in Math			
	Burbio	CSDH	R2L	EIPL	Burbio	CSDH	R2L	EIPL
	(1)	**(2)**	**(3)**	**(4)**	**(1)**	**(2)**	**(3)**	**(4)**
T_c	0.0006	0.0057*	0.0079**		-0.0007	0.0056**	0.0069**	
	(0.0023)	(0.0031)	(0.0036)		(0.0022)	(0.0027)	(0.0032)	
H_c	0.0009	0.0034	0.0048		-0.0004	0.0026	0.0034	
	(0.0029)	(0.0039)	(0.0036)		(0.0029)	(0.0035)	(0.0033)	
$EIPL_c$				0.0103***				0.0090***
				(0.0031)				(0.0029)
State FE	✔	✔	✔	✔	✔	✔	✔	✔
R-squared	0.34	0.35	0.35	0.35	0.33	0.35	0.34	0.34
# of counties	1,482	1,226	1,439	1,450	1,482	1,226	1,439	1,450

Notes: Each column reports coefficients from a weighted OLS regression with standard errors clustered at the county level in parentheses. Panel (a) shows estimates for online participation in Math, and panel (b) shows estimates for progress in Math. In each panel, columns (1), (2) and (3) use in-person and hybrid learning as measured by respectively Burbio, the COVID-19 school data hub (CSDH) and Return2Learn (R2L), and column (4) uses our EIPL measure. The regressions are estimated on average county-level in-person and hybrid learning or EIPL for the period from September 2020 to May 2021.

the strong negative correlation between T_c and H_c. In fact, according to the trackers, hybrid learning has no statistically significant impact on students' participation and progress in Math. This result is hard to interpret because, at the same time, two of the school trackers show that in-person learning has a positive effect, and hybrid learning is partly composed of in-person learning.

References

COVID-19 School Data Hub. School Learning Model Database, 2021. URL https://www.covidschooldatahub.com/for_researchers.

MIT Election Data and Science Lab. County Presidential Election Returns 2000-2020, 2018. URL https://doi.org/10.7910/DVN/VOQCHQ.

Erin M Fahle, Benjamin R Shear, Demetra Kalogrides, Sean F Reardon, Richard DiSalvo, and Andrew D Ho. Stanford Education Data Archive (Version 4.1), 2021. URL http://purl.stanford.edu/db586ns4974.

Nat Malkus. Federal COVID Elementary and Secondary School Emergency Relief funding district-level data compilation. Technical report, Return to Learn Tracker, American Enterprise Institute, 2021. URL http://www.returntolearntracker.net/esser/.

Education Week. Map: Where Has COVID-19 Closed Schools? Where Are They Open? (2020, July 28), 2021. URL https://www.edweek.org/leadership/map-where-are-schools-closed/2020/07.

Supplementary file for:
School Closures and Effective In-Person Learning during COVID-19

E Construction of Safegraph visit sample

This section presents additional information about the Safegraph visit data and sample used to create the EIPL dataset.

E.1 Matching of Safegraph POIs with NCES data

Our algorithm to match Safegraph's schools to the NCES's CCD and PSS files works as follows:

1. Prior to matching schools data to Safegraph, we deduplicate and pre-treat the Safegraph data by cleaning POIs' names and addresses. For names, we convert the capital letters to lower case and remove all the "%", "&", etc., numbers (if any), and spaces from the raw Safegraph location names. More importantly, we replace abbreviated school information in the Safegraph names by a complete descriptor using the following rules:[1]

Portion of the raw Safegraph name:	**Recoded as:**
elemsch	elementaryschool
highsch	highschool
kindergsch	kindergarten
middlesch	middleschool
primarysch	primarychool
schoolthe	school

   Last, we clean schools' addresses by using Stata's `stnd_address` command to standardize street address names.

2. We clean names and addresses in a similar way in the NCES's CCD and PSS files, where we have information on school names and addresses that describe the physical location of schools (street address and postal code). We clean school names by converting the capital letters to lower case and removing all the "%", "&", etc., numbers (if any), and spaces. We use Stata's `stnd_address` command to standardize street addresses.

3. We pool the cleaned CCD and PSS files, and then match them to Safegraph by applying the following consecutive attempts:

   (a) Attempt to directly merge schools sequentially in this order: (i) merge by name/address/zip-code, (ii) merge by name/zip-code;

   (b) Attempt to match schools based on GPS coordinates and school names. Within each local geographic area (defined by latitude × longitude rounded to the first decimal place), we measure (i) the geographic distance between schools based on GPS coordinates and (ii) the Levenshtein distance between school names (normalized by the length of the longest string of school name). We match schools that are closest to each other, provided that they are less than 250 meters away and that the Levenshtein distance is under 0.250.

[1] As an example, consider the Safegraph POI called "Big Spring Lake Kinderg Sch". After removing the spaces and converting the capital letters to lower case, we obtain "bigspringlakekindergsch". We then rename it as: "bigspringlakekindergarten". This enables us to increase the quality of the match to NCES data where typically the word "Kindergarten" is not abbreviated.

(c) Attempt to fuzzy-name match schools within each 5-digit zip codes sequentially in this order: (i) match on name/address, (ii) match on name, (iii) match on address. For fuzzy-name matching, we use Stata's `reclink2` command and define as high-quality matches those with a matching score higher than 0.85.[2,3]

E.2 Normalization and sample selection

An important concern when working with Safegraph's data is that changes in visits counts over time can be driven by changes in the sample of cell phone devices that Safegraph uses. Following large variations in the first two quarters of 2018, the sample size expands until mid-2019, then drops during the second half of 2019 and expands again in January of 2020. More importantly, the sample sizes drops substantially at the beginning of the pandemic and never recovers afterwards; in 2021 the sample size actually decreases relative to the second half of 2020.

Figure E1 illustrates the impact of these variations on counts of visits to all Safegraph's POIs with NAICS code 611110. In the upper panel, there is a clear upward trend in raw visits throughout 2018, 2019, and early 2020, as well as an incomplete recovery of visits in 2021 relative to pre-pandemic levels of visits. The bottom panel shows that normalizing by county-level counts of cell phone devices removes the trend in 2018 and 2020, while inducing visits at the end of 2019 and at the beginning of 2020 to be higher than before the Summer of 2019. The effects of normalization is also important for the recovery in 2021: normalized school visits return to their pre-pandemic levels, whereas in the not normalized data they remain about 25% lower. Motivated by these observations, throughout our analysis we normalize school visits with the weekly county-level counts of Safegraph cell phone devices.

In an effort to reduce measurement error, we implement the following sample restrictions:

- First, we drop schools where the raw visits count on average during the base period is less than 10, and schools where $\Delta\widetilde{v}_{j,t}$ is larger than 50 more than once during the based period. The goal of these first two restrictions is to ensure that the measurement of school visits for the base period are reliable enough to compare them with school visits in any other period. Together these restrictions reduce the sample size by 20%.

- Then, we drop schools where $\Delta\widetilde{v}_{j,t}$ is larger than 75 more than once, either during the period from beginning of September 2019 to November 2019 or the period from beginning of September 2020 to the end of the sample period. This procedure intends to purge the data from extreme values that affect the average of changes in visits in any given period. We use a larger threshold (75 instead of 50) to trim the data because it is to expected that the visits time series for each school are more volatile outside of the November 2019 to February 2020 period. This sample restriction reduces the sample size by an additional 10%.

The resulting "in-scope" dataset contains 73,194 schools or about two thirds of all schools that we successfully match to the CCD-PSS file. As a final data preparation step, we adjust changes in school visits $\Delta\widetilde{v}_{j,t}$ in the following way. First, we top-code $\Delta\widetilde{v}_{j,t}$ at 100%. Second, if in any week t outside of the reference period $\Delta\widetilde{v}_{j,t} > 25\%$ while $\Delta\widetilde{v}_{j,t-1} \leq 25\%$ and $\Delta\widetilde{v}_{j,t+1} \leq 25\%$, we replace $\Delta\widetilde{v}_{j,t}$ by the average of $\Delta\widetilde{v}_{j,t-1}$ and $\Delta\widetilde{v}_{j,t+1}$.[4]

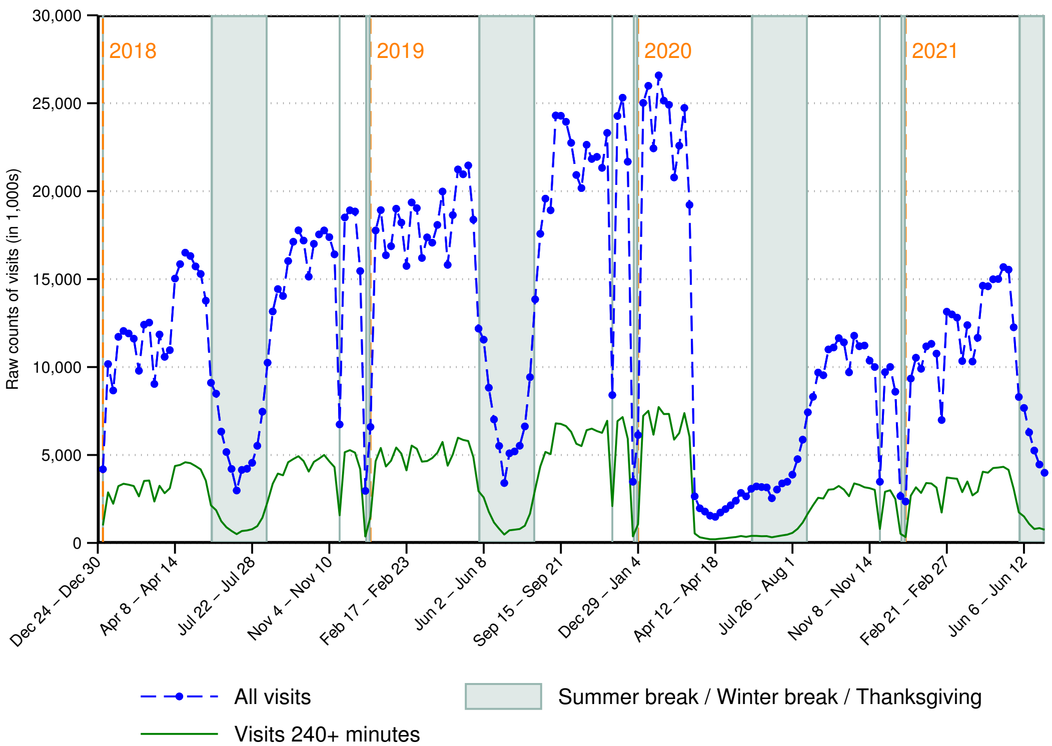
[2] To take one example of fuzzy name matches of high quality, consider Safegraph's "Big Spring Lake Kinderg Sch" described in Footnote 1. The name of this school in the CCD file is "Albertville Kindergarten and PreK". Through our algorithm, we obtain a fuzzy match at the name/address level (within the same 5-digit zip code) because the street addresses in Safegraph and in the CCD file turn out to be exactly the same. This, together with our update of the Safegraph's school name (step 1 of the matching algorithm), yields a matching score of 0.92 according to `reclink2` standard score metric.

[3] We manually compare a random sample of the matched schools to confirm that the thresholds (250 meters for the geographic distance, 0.250 for the Levenshtein distance, 0.85 for Stata's `reclink2` match score) are good markers of high- vs low-quality matches.

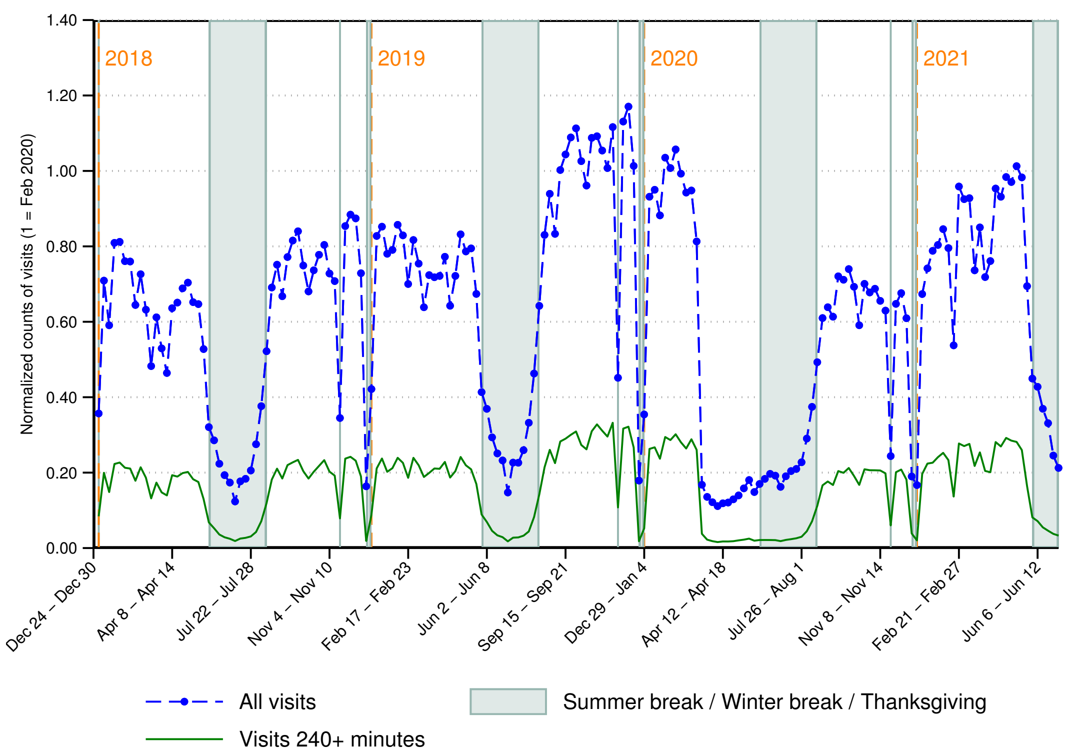
[4] This adjustment implements the assumption that schools did not reopen for only one week at a time.

Figure E1: Safegraph: Aggregate time series of school visits

(a) Raw counts of visits



(b) Normalized counts of visits



Notes: The figures show the raw (upper panel) and normalized (lower panel) counts of total weekly visits and counts of visits longer than 240 minutes to all Safegraph POI with NAICS code 611110 ("Elementary and secondary schools").

E.3 School weights

We augment the dataset with school-level weights to alleviate concerns about its representativeness after filtering out schools with sparse or noisy visit data. We estimate a Probit model where the left-hand side variable is an indicator y_j that takes the value of 1 if school j is included in the dataset of our analysis and is 0 otherwise. The regressors of the Probit model are: county-level shares of married adults, county-level shares of High School and College workers, a cubic polynomial of county population, population density, dummy variables for local area types of the school (i.e., city, suburban, town or rural area) and dummy variables for the nine U.S. Census divisions. Then, we weight each public school by the inverse of the predicted probability $\hat{\Pr}\{y_j = 1\}$, and each private school by its PSS sampling weight times the inverse of the predicted probability $\hat{\Pr}\{y_j = 1\}$.[5] We check the quality of this adjustment by comparing the weighted counts of students, teachers, and schools in the data to the same counts based on the pooled CCD/PSS file (i.e. those reported in the second column of the Table "Comparison with the NCES digest of education's statistics" of the Online appendix).

E.4 A closer look at changes in school visits

It should be noticed that even after implement sample restrictions to exclude POIs with sparse or noisy visits data, there remains substantial variations in changes in school visits. For instance a non-trivial share of changes in visits, $\Delta\widetilde{v}_{j,t}$, fall outside of the [-20%, +20%] interval in September and October of 2019 (that is, before the base period) and in January and February of 2020 (during the base period) for schools that we retain in our analysis. Some of this dispersion, however, may capture "true" variations in school activity across months. For example, some schools may not reopen right in the beginning of September 2019, and similarly, in January 2020, some schools might start later than others.

Figure E2 plots the distribution of average $\Delta\widetilde{v}_{j,t}$ during different months of the COVID-19 pandemic. The upper panel focuses on the first 4 months of the pandemic. The top left plot in this panel shows that $\Delta\widetilde{v}_{j,t}$ does well in capturing week-to-week variations: most schools were open during at least the first two weeks of March 2020 before being shut down, and as a result the change in school visits averaged over the 4 weeks of this month is -46 on average. In the other plots of the upper panel, the shift closer to -100% is obviously indicative of school closures.[6]

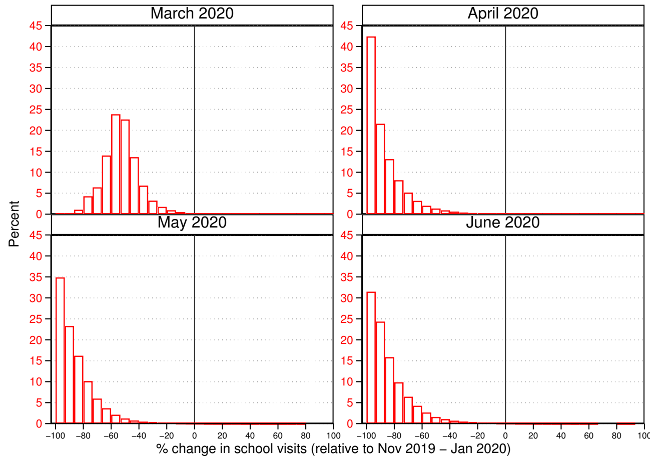
The middle and lower panels of Figure E2 show the distribution of average $\Delta\widetilde{v}_{j,t}$ during the Fall of 2020 and Spring of 2021. Note that the scale on the vertical axes of the plots is the same in the two panels. In panel (b), we see a recovery of $\Delta\widetilde{v}_{j,t}$ relative to the first few months of the pandemic, which is likely indicative of school reopenings in some regions. Then, we observe a slight reversal in November and December relative to September-October 2020, which is possibly linked to a tightening of health restrictions but also due to the fact that both months include one week of vacation (Thanksgiving in November 2020, Christmas in December 2020). The lower panel of Figure E2 shows a clearer recovery in school visits, though with substantial mass around 0 or higher.

[5] Since the CCD contains the universe of public schools, the sampling weight of public schools is 1 and therefore the adjusted weight is 1 divided by the probability of selection into the "in scope" dataset. Across all schools, the final weights that we obtain range from 1.16 to 122.3 with an average of 1.64 and a median of 1.51. For public schools, the weights range from 1.16 to 7.47 with an average of 1.56 and a median of 1.47. The larger weights of the "in scope" dataset are for private schools, but the large values come from the PSS sampling weights (which can go all the way up to a value of 75), as opposed to reflecting very small values of $\hat{\Pr}\{y_j = 1\}$.

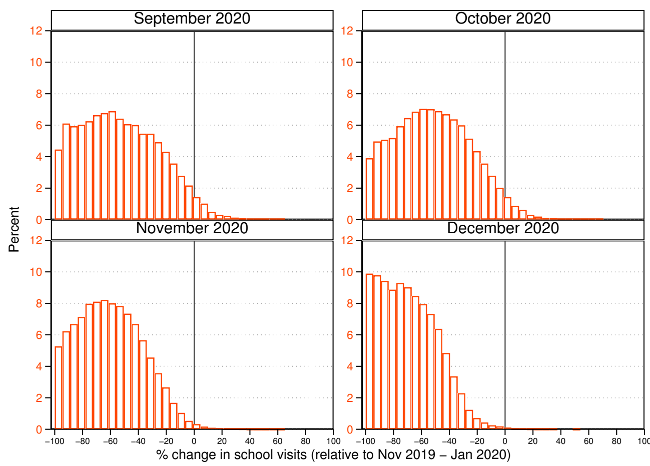
[6] In April, May and June in the upper panel of Figure E2, we observe some schools with changes in visits not lower than -60% or -80%. To understand how this relates to the upper map of county-level loss of EIPL (Figure 1, main text), recall that: 1) these changes in visits are translated into EIPL by being multiplied by a coefficient that can be greater than 1 as shown in Table 1 of the main text, and 2) week-to-week variations in visits at the individual school level imply that a school might have $\Delta\widetilde{v}_{j,t}$ between, say, -60% and -80% in May and between -80% and -100% in April and June. The latter source of variation is not present in Figure 1 since the data is averaged over longer periods of time.

Figure E2: Distribution of changes in school visits during the pandemic

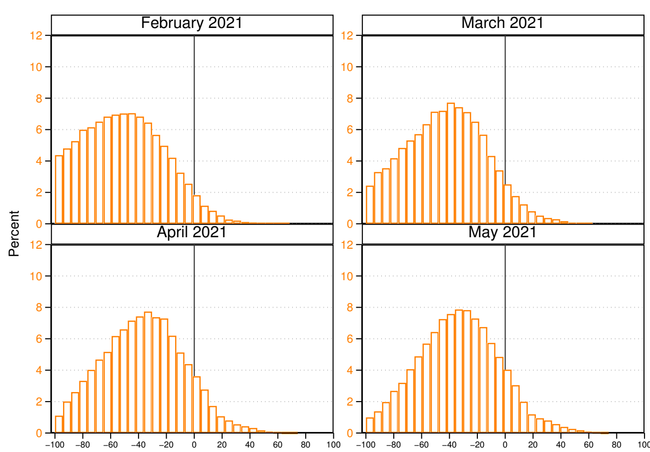
(a) March 2020 to May 2020



(b) September 2020 to December 2020



(c) January 2021 to May 2021



Notes: The figures show the distribution of the average change in school visits at points in time during the pandemic.

F Relation of schooling mode trackers with EIPL

It is instructive to compare EIPL with the schooling mode trackers. In Table F1, we merge each tracker with our EIPL data (aggregated to either the district, county, or state level i, depending on the tracker) and calculate the bivariate correlation between average EIPL and the average share of in-person (T_i), hybrid (H_i) and remote (R_i) learning, respectively, for the time period covered by each of the schooling mode trackers.

Table F1: Effective In-Person Learning compared to schooling mode trackers

	Burbio	CRPE	CSDH school	CSDH district	EdWeek	ESOS	IES-SSD	MCH	R2L
	(1)	(2)	(3a)	(3b)	(4)	(5)	(6)	(7)	(8)
T_i	73.3	78.8	58.9	73.1	92.6	72.8	89.3	66.6	63.3
	(1.25)	(2.95)	(0.41)	(1.27)	(8.03)	(0.75)	(8.07)	(0.93)	(0.87)
H_i	-20.2	0.50	-14.2	2.4	-27.3	-28.9	-11.0	-15.3	-11.2
	(1.80)	(4.79)	(0.50)	(1.86)	(20.5)	(1.04)	(15.9)	(0.96)	(1.11)
R_i	-73.6	-69.6	-61.7	-74.8	-80.7	-72.1	-82.0	-60.7	-70.6
	(1.25)	(3.44)	(0.40)	(1.24)	(12.6)	(0.75)	(10.06)	(0.88)	(0.79)
# of geo. units	2,953 counties	438 districts	39,629 schools	10,275 districts	51 states	8,497 districts	51 states	11,991 districts	7,953 districts
# of weeks	45	3	49	49	43	2	5	2	45
% of data covered	94.0	91.8	66.2	73.3	100	92.4	100	71.5	92.4

Notes: The table reports the correlation between EIPL and the average share of in-person (T_i), hybrid (H_i) and remote (R_i) learning provided in the schooling mode trackers from: Burbio, the Center on reinventing public education (CRPE), the COVID-19 school data hub (CSDH), the Elementary school operating status (ESOS) database, the School survey dashboard of the Institute of Education Sciences (IES-SSD), MCH strategic data (MCH), and Return2Learn (R2L). EIPL, in-person, hybrid and remote learning are averaged over the weeks covered by the school tracker in each column of the table. Standard errors are in parentheses. The lower panel reports the number of overlapping geographic units and weeks, and fraction of the school tracker data that are covered by the EIPL database.

Across all trackers, the EIPL measure is strongly positively (negatively) correlated with the share of in-person (remote) learning. It is also remarkable that, despite the weak correlations across pairs of trackers evidenced in Figure 1 in the paper, the correlations with EIPL are similar, suggesting that EIPL captures a common component. In contrast, the correlation with the share of hybrid learning is small and in some cases insignificantly different from zero. This reflects a fundamental characteristic of hybrid learning, that its relation with EIPL is non-linear. For regions that chose to keep schools closed for much of the 2020-21 school year, hybrid learning is low and so is EIPL. For regions that chose to reopen schools for most of the year, hybrid learning is also low but, naturally, EIPL is high. For regions in-between, hybrid learning is high while EIPL is moderate. This inverse hump-shaped pattern of hybrid learning with respect to EIPL represents an important, though perhaps underappreciated challenge for empirical analyses that use the share of hybrid learning as a regression variable. By combining in-person and hybrid learning, our estimation of EIPL circumvents this issue.

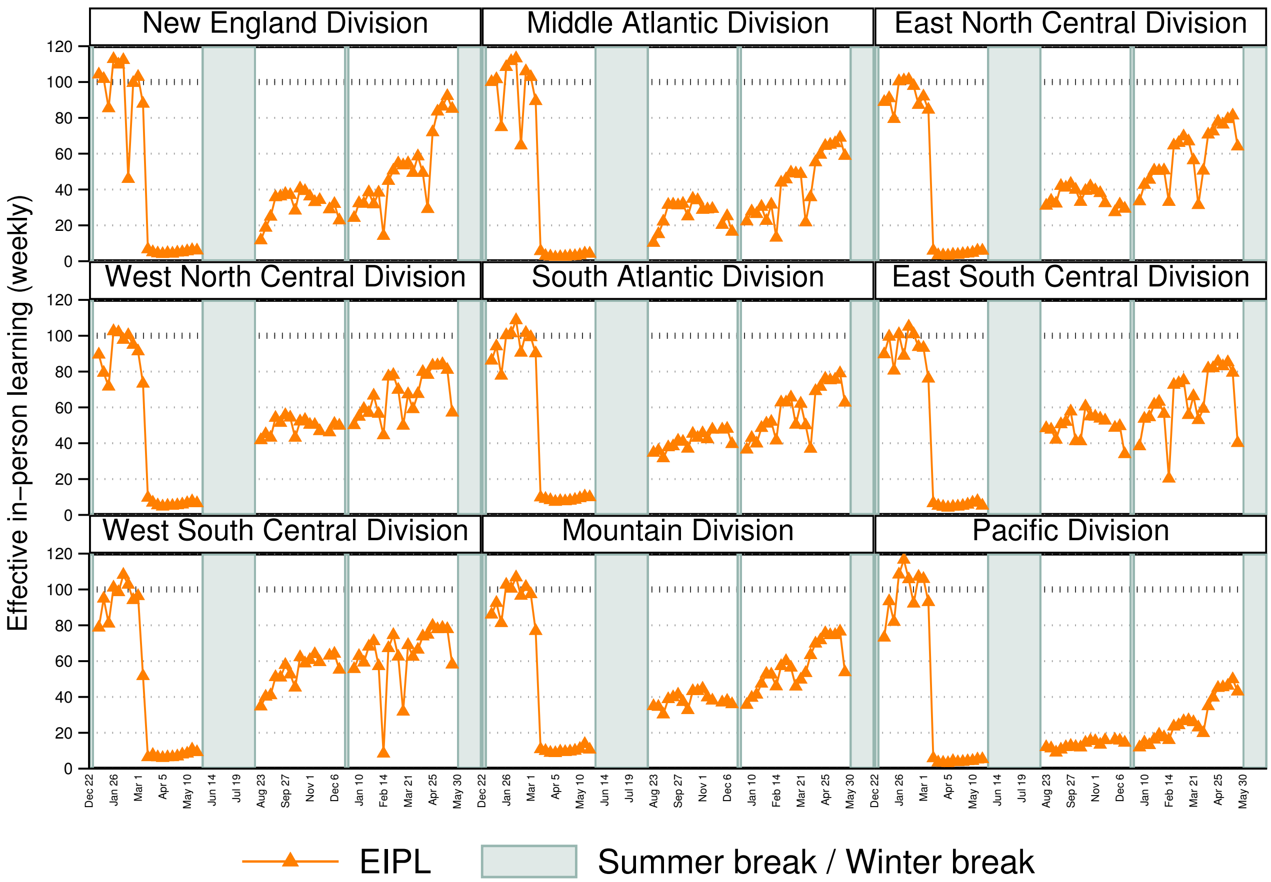
A final interesting result from Table F1 is that the EIPL dataset includes a large fraction of the schools covered across the different trackers. Two thirds of the schools in CSDH are present in our dataset; 70 to 90 percent of the school districts covered in other trackers are included in ours; and almost 95 percent of the counties from Burbio cover schools that are in our dataset. While the overlap is important, the EIPL dataset has clear advantages through its granularity: it allows to study schools separately by type (public charter, public non-charter, private religious, private nonreligious) and grade (elementary school, middle school, high school, or a combination thereof), with each school equipped with a sampling weight to ensure representativeness. The school-level EIPL measure aggregates up in a way that is consistent with the categorical indicators of the three learning modes available from other, more aggregated data, while being easier to work with as it is a continuous variable.

G Additional tables and figures

G.1 Regional disparities in EIPL over time

Figure G1 summarizes the temporal and geographic variation in EIPL by averaging weekly student-weighted EIPL for each of the nine U.S. Census Divisions. While EIPL drops to near zero for all divisions between March and May 2020, we see large differences during the 2020-21 school year. EIPL in states in the West North Central, East South Central and West South Central division quickly increase to 60% from September 2020 through December 2020 and climb to over 80% from January through May 2021. In contrast, EIPL in states in the New England, the Middle Atlantic and especially the Pacific division remains below 50% for most of the 2020-21 school year. Some of the regional disparities appear right after the end of the Summer break, while others built up later during the school year.

Figure G1: Weekly effective in-person learning, by Census divisions

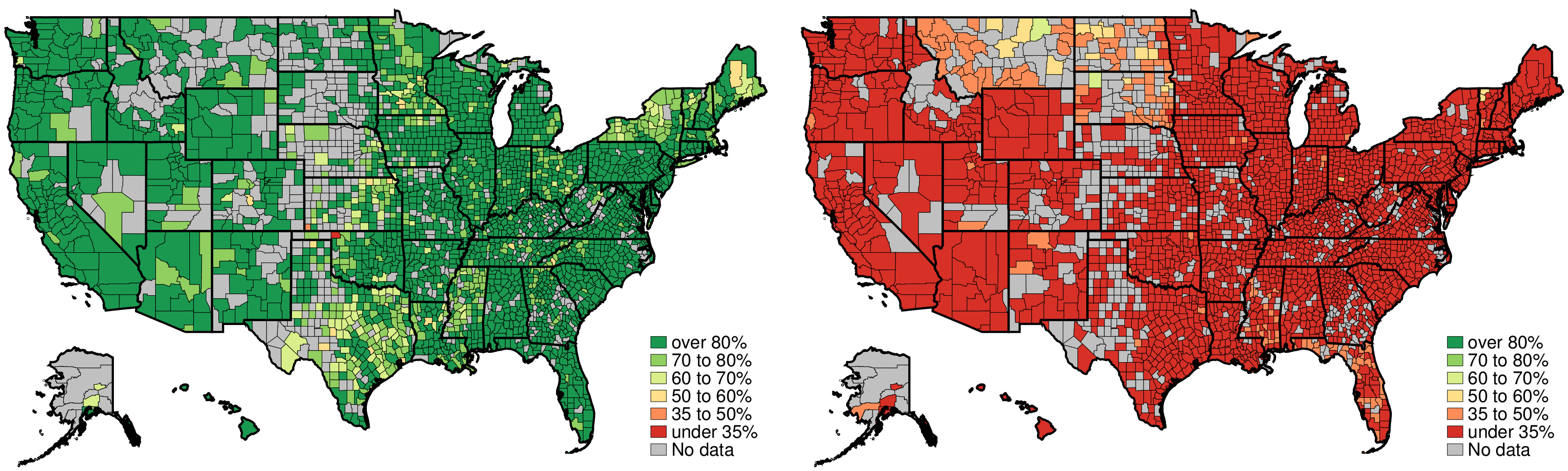


Notes: The figure shows student-weighted, weekly effective in-person learning for the different U.S. Census Divisions: New England (CT, MA, ME, NH, RI, VT), Middle Atlantic (NY, NJ, PA), East North Central (IL, IN, MI, OH, WI), West North Central (IA, KS, MN, MO, NE, ND, SD), South Atlantic (DE, FL, GA, MD, NC, SC, VA, WV), East South Central (AL, KY, MS, TN), West South Central (AR, LA, OK, TX), Mountain (AZ, CO, ID, MT, NM, NV, UT, WY), Pacific (CA, OR, WA).

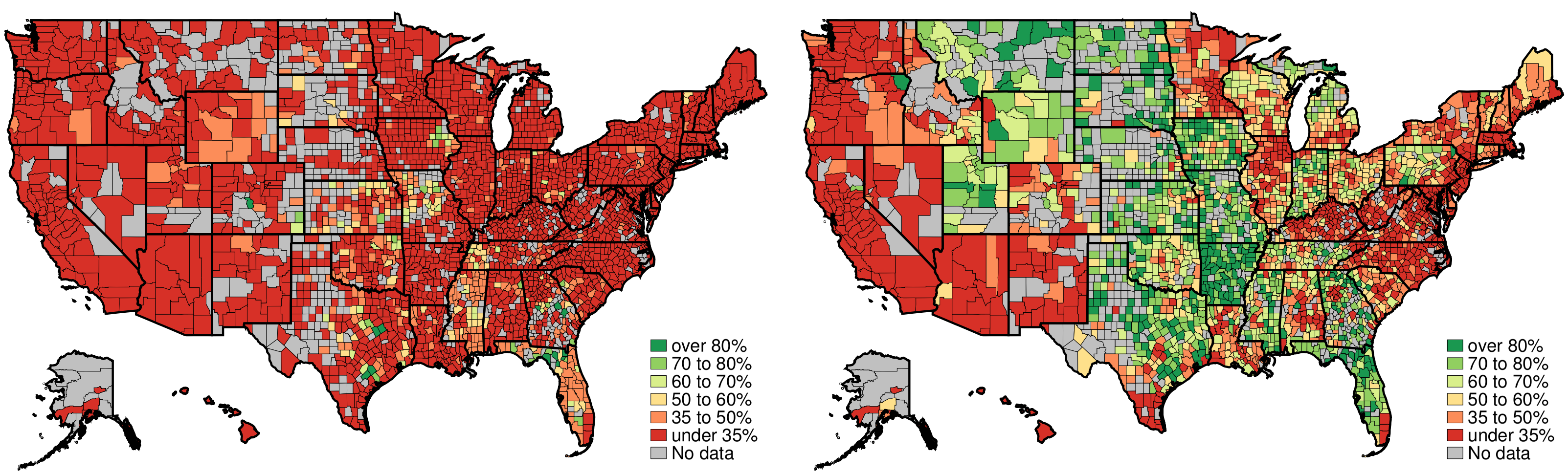
The upper panel of Figure G2 reinforces the finding from Figure G1, that EIPL at the beginning of the pandemic drops sharply across all regions of the country. The left plot in this panel shows county-level EIPL on average from mid-February 2020 until mid-March 2020. It shows levels of EIPL over 80% in

Figure G2: Effective in-person learning across U.S. counties

(a) 4 weeks before (left) and after (right) March 15, 2020

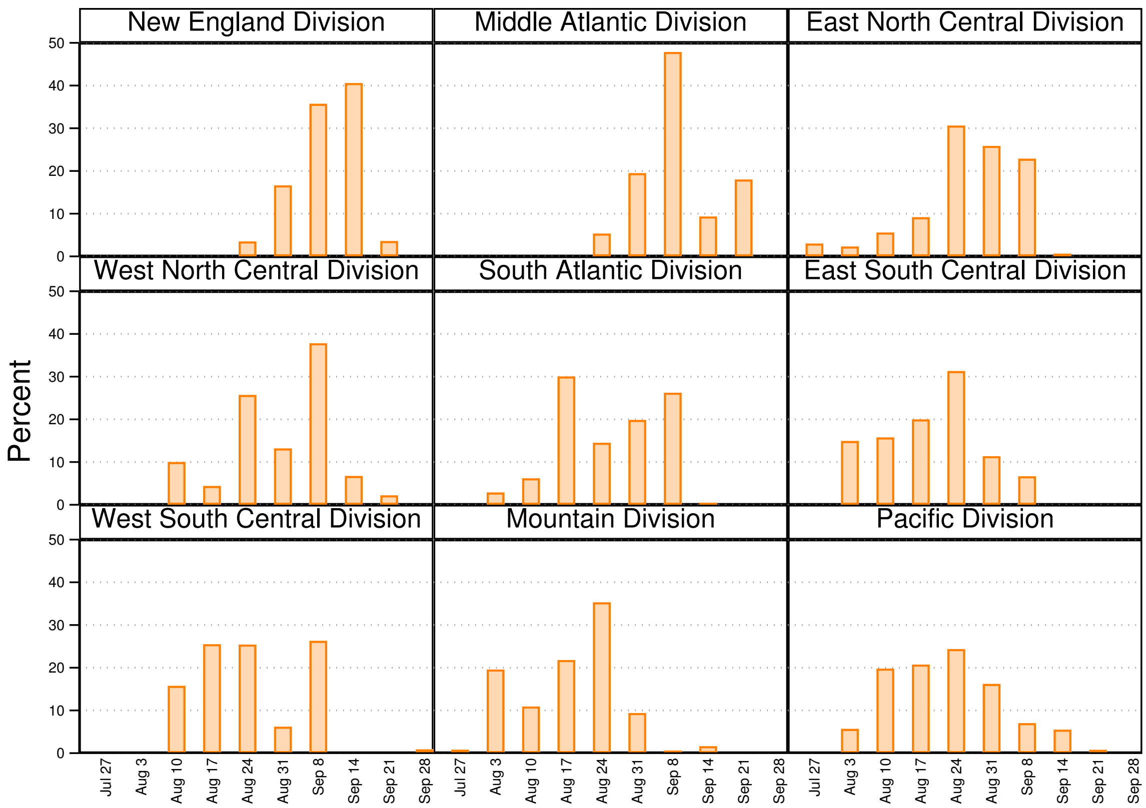


(b) 4 weeks before (left) and after (right) back to school dates according to pre-COVID school district calendars



Notes: The figure shows the student-weighted average county EIPL for all counties for which we have reliable data on at least three schools.

Figure G3: Back to school dates (based on pre-COVID school district calendars), by Census divisions



Notes: The figure shows the percent of public schools that would open on the days in question according to pre-COVID school district calendars in each Census division: New England (CT, MA, ME, NH, RI, VT), Middle Atlantic (NY, NJ, PA), East North Central (IL, IN, MI, OH, WI), West North Central (IA, KS, MN, MO, NE, ND, SD), South Atlantic (DE, FL, GA, MD, NC, SC, VA, WV), East South Central (AL, KY, MS, TN), West South Central (AR, LA, OK, TX), Mountain (AZ, CO, ID, MT, NM, NV, UT, WY), Pacific (CA, OR, WA).

almost every county. By contrast, the plot on the right-hand-side of this panel showing EIPL on average from mid-March 2020 to mid-April 2020 is almost uniformly "red". There are a few exceptions, notably in Montana, North and South Dakota, which seem to line up with state-specific responses in the adoption of mitigation strategies at the beginning of the COVID pandemic. For example, the governor of the state of South Dakota adopted an executive order to encourage social distancing and remote work in mid-March of 2020 but resisted imposing a mandatory, state-wide lockdown, and later on ruled out a state mandate on the wearing of face masks in public spaces.

To further analyze the regional and temporal disparities in EIPL, we next look at EIPL around the time when schools *usually* reopen after the Summer break. To this end, we use information from Burbio about the week when most public schools within a county usually reopen. This information, which is reported in Figure G3 by showing back-to-school dates for each of the nine Census divisions, reveals major differences across different parts of the country. For example, in the East South and West South Central divisions, students usually head back to school at the beginning of August, whereas in New England and in the Middle Atlantic division back-to-school dates are typically after Labor Day.[7,8] We use this information (available at the county-level) in the lower panel of Figure G2 to check EIPL in the four weeks before and after back-to-school dates. As can be seen on the left-hand side of the figure, our measure of EIPL indicates close-to-zero in-person learning for most counties before their usual back-to-school dates. That the picture is not uniformly red could reflect inaccuracies of our EIPL measure, but may also be explained by (i) usual back-to-school dates which are not uniform within a county or not well measured by Burbio, (ii) reopening of private schools that may be asynchronous with that of public schools, (iii) schools reopening for in-person learning earlier than what they usually do in normal times. Then, on the right-hand side of the panel, we observe many counties shifting from close-to-zero to much higher levels of EIPL. On the other hand, and as expected, most counties where we measure low EIPL throughout the school year 2020-21 are in red color in the lower panel of Figure G2 before and after usual back-to-school dates.

G.2 Additional regression results

This section presents results from several regressions summarized in Section 4 of the paper.

G.2.1 Quasi-univariate relations. We begin in Table G1 with results of regressing EIPL separately on each of the three affluence measures together with the share of non-white students and controls for school type and school grade. As mentioned, when controlling for a school's share of non-white students does not change the coefficients on household income and education substantially, but it turns the coefficient on the share of dual-headed households to negative. Also, adding controls for school type and school grade does not change the results noticeably, and the estimates on these controls are in line with the univariate relations between EIPL and schools' type and grade. As mentioned in the text, the reason we do not include all three affluence measures together in the regressions is that the high correlation between them make it difficult to interpret the estimates.

G.2.2 Role of school variables. Next, we focus on the role of (pre-COVID) test scores, school size, and school funding. Except for school enrollment, the different variables are not available for private schools. We therefore focus here on public schools. Column (1) in Table G2 repeats the results from

[7] Research from the Pew Research Center indicates that these differences are historically related to preferences over teenagers taking on work summer jobs, constraints that limit the time when families can take vacations, and the economic importance of tourism and hospitality industries.

[8] In Figure G1, the shaded area denoting the Summer break of 2020 covers the weeks from May 31st until August 22nd. For many schools across the country, this time interval is only an approximation of the Summer break since, as shown in Figure G3, back-to-school dates are not uniform across regions.

Table G1: The inverse relationship of effective in-person learning with affluence and race

Dependent variable	Effective in-person learning (EIPL)					
	(a) Public schools			**(b)** Private schools		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Zip-level household income	-5.29***			-5.32***		
	(0.46)			(0.53)		
Zip-level share of college educated		-7.74***			-8.63***	
		(0.43)			(0.75)	
Zip-level share of dual-headed households			-3.05***			-3.31***
			(0.70)			(0.68)
School share of non-white students	-22.93***	-22.57***	-23.95***	-8.72***	-9.01***	-9.12***
	(1.21)	(1.25)	(1.43)	(0.77)	(0.80)	(0.83)
School type and grade controls	✔	✔	✔	✔	✔	✔
R-squared	0.12	0.13	0.11	0.07	0.07	0.05
# of counties	2,951	2,951	2,951	1,444	1,444	1,444
# of schools	60,054	60,054	60,054	9,651	9,651	9,650
# of school districts	12,505	12,505	12,505			

Notes: Each column reports coefficients from a weighted OLS regression with standard errors clustered at the county level in parentheses and school weights calculated as explained in Subsection E.3. The regressions are estimated on average school EIPL for the period from September 2020 to May 2021. Panel (a) shows estimates for the public school sample, and panel (b) shows estimates for the private school sample. The school type fixed effects consists of indicators for charter school and non-charter school for the public school sample, and religious school and non-religious school for the private school sample. The school grade fixed effects consist of indicators for elementary vs. middle vs. high. vs. combined school for both samples.

Table G1 above as a reference[9]; Column (2) adds district-level test scores to the regression; Column (3) adds school size; Column (4) adds school spending and ESSER funding; and Column (5) adds the four variables jointly. The table shows that controlling for the school variables separately or jointly has no impact on the results.

G.2.3 Role of geography. Table G3 analyzes the role of geography. Column (1) in panel (a) repeats the final regression in Table G2 above for reference. Columns (2) and (3) add fixed effects for the state, respectively the county in which the school is located. The consequences of controlling for these more detailed geographical effects are important, raising the explanatory power of the regressions to almost one third, and can be summarized as follows.

First, the inverse relation between EIPL and local affluence is cut in half when the state fixed effect is added, and essentially disappears when the county fixed effect is added. Similarly, the association of EIPL with local education is substantially reduced although it remains negative, implying that even within counties, schools located in zip-codes with a higher share of college-educated households provided on average somewhat lower EIPL. We conclude from these estimates that EIPL is negatively related to affluence and education primarily because less affluent and less educated areas of the county have public schools that provided more EIPL during the 2020-21 school year. Second and contrary to affluence and education, the inverse relation between EIPL and the share of non-white students is unaffected by state fixed effects and is reduced by only about one third by the county fixed effect. So, even within counties and controlling for affluence, education and other school characteristics, there are clear racial differences in that schools with a larger share of non-white students provided on average substantially lower EIPL. Third, the negative coefficient estimate on school size remains unaffected by the state and county fixed

[9] The estimates for household income, share of college educated and school neighborhood poverty are exactly as in columns (1) - (3) of Table G1. The estimate for share of non-white students is slightly different because this estimate is obtained while controlling jointly for all three measures.

Table G2: The role of test scores, school size, and school funding

Dependent variable	Effective in-person learning (EIPL)				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Zip-level household income[a]	-5.29***	-6.57***	-4.63***	-5.58***	-5.27***
	(0.46)	(0.54)	(0.46)	(0.54)	(0.58)
Zip-level share of College educated[a]	-7.74***	-9.99***	-6.97***	-7.97***	-8.51***
	(0.43)	(0.40)	(0.44)	(0.45)	(0.43)
Zip-level share of dual-headed households[a]	-3.05***	-2.61***	-2.15***	-3.67***	-2.92***
	(0.70)	(0.79)	(0.70)	(0.73)	(0.74)
School share of non-white students[b]	-21.14***	-18.62***	-19.79***	-19.40***	-15.27***
	(1.60)	(1.48)	(1.55)	(1.43)	(1.34)
Mean maths-RLA test scores[b]		4.48***			5.12***
		(0.68)			(0.81)
Student enrollment[b]			-2.43***		-3.56***
			(0.28)		(0.28)
School spending per student[b]				-3.60***	-4.96***
				(0.63)	(0.68)
ESSER funding per student[b]				-2.19***	-0.52
				(0.66)	(0.76)
School type and grade controls	✔	✔	✔	✔	✔
R-squared	0.13	0.14	0.14	0.14	0.15
# of counties	2,951	2,951	2,951	2,863	2,863
# of schools	60,054	60,054	60,054	56,632	56,632
# of school districts	12,505	12,505	12,505	11,391	11,391

Notes: Each column reports coefficients from a weighted OLS regression on the public school sample, with standard errors clustered at the county level in parentheses and school weights calculated as explained in Subsection E.3. The regressions are estimated on average school EIPL for the period from September 2020 to May 2021. The school type fixed effects consists of indicators for charter school and non-charter school, and the school grade fixed effects consist of indicators for elementary vs. middle vs. high. vs. combined school for both samples. The coefficient estimates for the affluence measures, denoted by [a], are the result of separate regressions with each one of the measures in combination with the other variables below. The coefficient estimates for the other regressors denoted by [b] are the result of regressions where the three affluence measures are included jointly.

Table G3: The importance of geography

Dependent variable	Effective in-person learning (EIPL)					
	(a) Public schools			**(b)** Private schools		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Zip-level household income[a]	-5.27***	-2.86***	-0.54***	-5.05***	-2.74***	-0.49
	(0.58)	(0.34)	(0.18)	(0.51)	(0.38)	(0.41)
Zip-level share of college educated[a]	-8.51***	-6.47***	-2.59***	-8.26***	-5.96***	-2.13***
	(0.43)	(0.37)	(0.23)	(0.74)	(0.58)	(0.61)
Zip-level share of dual-headed households [a]	-2.92***	-0.66*	0.19	-3.15***	-0.34	1.19***
	(0.74)	(0.35)	(0.22)	(0.66)	(0.44)	(0.43)
School share of non-white students[b]	-15.27***	-16.24***	-7.07***	-8.49***	-7.10***	-3.67***
	(1.34)	(0.93)	(0.59)	(0.89)	(0.66)	(0.63)
Mean maths-RLA test scores[b]	5.12***	3.11***	3.67***			
	(0.81)	(0.60)	(0.41)			
Student enrollment[b]	-3.56***	-3.38***	-2.97***	-1.03***	-1.19***	-0.76***
	(0.28)	(0.22)	(0.21)	(0.37)	(0.29)	(0.27)
School spending per student[b]	-4.96***	0.07	0.56***			
	(0.68)	(0.29)	(0.21)			
ESSER funding per student[b]	-0.52	-0.78	-1.17***			
	(0.76)	(0.51)	(0.42)			
School type and grade controls	✔	✔	✔	✔	✔	✔
State FE		✔			✔	
County FE			✔			✔
R-squared	0.15	0.27	0.35	0.08	0.14	0.24
# of counties	2,863	2,863	2,863	1,444	1,444	1,444
# of schools	56,632	56,632	56,632	9,650	9,650	9,650
# of school districts	11,391	11,391	11,391			

Notes: Each column reports coefficients from a weighted OLS regression on the public (panel (a)) and private (panel (b)) school samples, with standard errors clustered at the county level in parentheses and school weights calculated as explained in Subsection E.3. The regressions are estimated on average school EIPL for the period from September 2020 to May 2021. The school type fixed effects consists of indicators for charter school and non-charter school, and the school grade fixed effects consist of indicators for elementary vs. middle vs. high. vs. combined school for both samples. The coefficient estimates for the affluence measures, denoted by ^{a}, are the results of separate regressions with each one of the measures in combination with the other variables below. The coefficient estimates for the other regressors denoted by ^{b} are the result of regressions where the three affluence measures are included jointly.

effect. The result is interesting because it suggests that smaller schools reopened to in-person learning more quickly than larger schools, perhaps because the logistical challenges of reopening or equity concerns about reopening only certain grades were less important.[10]

Panel (b) of Table G3 repeats the same set of regressions for private schools. As with our analysis of public schools, we find that the role of affluence is substantially reduced (the coefficient on household income becomes not statistically different from zero), and the coefficient on the share of non-white students decreases too. The fixed effects raises the explanatory power of the regressions to over 20%. In sum, panel (b) supports the conclusion that the negative relation between EIPL and affluence and education is driven by less affluent and less educated areas of the county having schools (public, but also private) that provide more EIPL during the 2020-21 school year. Last, the regressions show that our findings on the negative role of school size extend to private schools, but that the effect is quantitatively smaller.

G.2.4 Share of nonwhite students. Our results show that the share of non-white students among the schools' body is associated with lower EIPL during 2020-21. In Table G4, we present additional results regarding this finding by replacing the share of non-white students by shares of students of different races or ethnicities. Columns (1), (3) and (5) of panel (a) repeat the results from the main regressions discussed in the text (Figure 4), while columns (2), (4) and (6) show the estimates for the different races/ethnicities; Panel (b) performs a similar analysis for private schools. Table G4 shows that the coefficient on the school share of non-white students is mainly driven by the share of Hispanic students, and to a much lesser extent by that of Black students. We hypothesize that the important role of the share of Hispanic students is related to the ethnic makeup of schools in states of the South-western part of the country (California, New Mexico), where as shown in Figure 1 EIPL has remained very low throughout the school year of 2020-21.

G.2.5 District vs. school-level test scores. Our main regression uses pre-COVID test scores at the level of school districts for reasons of data availability. In Table G5, we show effects of using school-level test scores, which are available for about 33,000 schools in our dataset vs. 57,000 for the district-level test scores. Columns (1) and (2) repeat results from the main regressions discussed in the text (Figure 4). As can be seen the addition of county-level controls barely changes the coefficient on test scores. Then, in Columns (3) and (4), we restrict the regression to schools for which we also have school-level test scores available from SEDA. The coefficient on district-level test scores increases slightly. Finally in Columns (5) and (6), we replace district-level test scores by school-level test scores. The magnitude of the effects of test scores changes, but not by much. The effects of the addition of county-level controls is similar to that obtained under our main regression.

G.2.6 Other indicators of local affluence. Table G6 presents additional results of introducing other indicators of local affluence. The indicators considered are: the Opportunity Atlas (OA)'s measure of upwards mobility as measured by the mean household income rank for children whose parents were at the 25th percentile of the national income distribution, where incomes for children are measured as mean earnings in 2014-2015 when they were between the ages 31-37; OA's fraction of children born in 1978-1983 birth cohorts with parents at the 25th percentile of the national income distribution who were incarcerated in 2010; OA's measured average rent for two-bedroom apartments in 2015; and EDGE's school neighborhood poverty estimates.[11] Columns (1), (3), (5) and (7) run a quasi-univariate regression,

[10] As noted above, the regressions control for whether the school is an elementary school, high school, or combined school; but these controls are relatively coarse and there may substantial variations in the number of grades served by a school even within these categories.

[11] To construct school neighborhood poverty estimates, EDGE uses data from the Census Bureau's American Community Survey to compute local income-to-poverty ratios (IPR). IPRs measure the percentage of family income that is above or below the federal poverty threshold set for the family's size and structure. IPRs are then aggregated to the levels of the school neighborhood as identified by EDGE.

Table G4: Robustness check: racial makeup of schools

Dependent variable	Effective in-person learning (EIPL)											
	(a) Public schools						**(b) Private schools**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Non-white students	-20.29***		-15.27***		-6.18***		-8.46***		-8.49***		-3.96***	
	(1.60)		(1.34)		(0.86)		(0.89)		(0.89)		(0.61)	
Black students		-3.08***		-1.66***		-0.70**		-1.63***		-1.68***		-0.80***
		(0.35)		(0.36)		(0.31)		(0.21)		(0.21)		(0.19)
Hispanic students		-10.75***		-8.00***		-3.19***		-2.74***		-2.74***		-1.50***
		(0.89)		(0.69)		(0.54)		(0.47)		(0.47)		(0.24)
American indian students		-0.08***		-0.03		-0.10***		-0.00**		-0.00**		-0.00**
		(0.02)		(0.02)		(0.02)		(0.00)		(0.00)		(0.00)
Hawaiian students		-0.08***		-0.04**		0.07		-0.00***		-0.00***		-0.00**
		(0.02)		(0.02)		(0.09)		(0.00)		(0.00)		(0.00)
Asian students		-1.93***		-1.74***		-0.73***		-1.27***		-1.26***		-0.16
		(0.14)		(0.14)		(0.10)		(0.22)		(0.22)		(0.22)
Students with 2 or more races		-2.62***		-2.29***		-1.02***		-1.83***		-1.77***		-0.69*
		(0.43)		(0.41)		(0.31)		(0.44)		(0.43)		(0.38)
Local affluence controls	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
School type and grade controls	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Other school/district variables			✔	✔	✔	✔			✔	✔	✔	✔
County controls					✔	✔					✔	✔
R-squared	0.13	0.13	0.15	0.16	0.27	0.27	0.08	0.08	0.08	0.08	0.14	0.14
# of counties	2,863	2,863	2,863	2,863	2,633	2,633	1,444	1,444	1,444	1,444	1,444	1,444
# of schools	56,632	56,632	56,632	56,632	56,019	56,019	9,650	9,650	9,650	9,650	9,650	9,650
# of school districts	11,391	11,391	11,391	11,391	10,974	10,974						

Notes: Each column reports coefficients from a weighted OLS regression on the public (panel (a)) and private (panel (b)) school samples, with standard errors clustered at the county level in parentheses and school weights calculated as explained in Subsection E.3. The regressions are estimated on EIPL for the period from September 2020 to May 2021. Local affluence variables consist of zip-level mean household income, share of adults with College or higher education, share of dual-headed household with children. The school type fixed effects consists of indicators for charter school and non-charter school, and the school grade fixed effects consist of indicators for elementary vs. middle vs. high. vs. combined school for both samples. Other school/district variables consist of district-level test scores, school size, school spending per student and ESSER funding per student. County controls consist of pre-pandemic ICU bed capacity, two-week lagged county COVID case and death rates, population density in the county, dummies for rural-urban continuum codes, maximum weekly temperature in the county, and dummies for the various non-pharmaceutical interventions (see Appendix B.4).

Table G5: Robustness check: District-level vs. school-level test scores

Dependent variable	Effective in-person learning (EIPL)					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
District-level mean math-RLA test scores	4.88***	4.47***	5.69***	4.82***		
	(0.67)	(0.59)	(0.72)	(0.67)		
School-level mean math-RLA test scores					3.27***	2.34***
					(0.66)	(0.47)
Student enrollment		-2.78***		-4.37***		-4.37***
		(0.25)		(0.30)		(0.30)
School spending per student		-1.03***		-1.81***		-1.39***
		(0.32)		(0.40)		(0.40)
ESSER funding per student		-0.94		-0.83		-2.62***
		(0.62)		(0.70)		(0.69)
School-level test scores available			✔	✔	✔	✔
Local affluence	✔	✔	✔	✔	✔	✔
School type and grade controls	✔	✔	✔	✔	✔	✔
County controls		✔		✔		✔
R-squared	0.13	0.27	0.11	0.25	0.11	0.25
# of schools	56,632	56,632	32,729	32,729	32,729	32,729
# of school districts	11,391	11,391	8,166	8,166	8,166	8,166
# of counties	2,863	2,863	2,612	2,612	2,612	2,612

Notes: Each column reports coefficients from a weighted OLS regression on the public school sample, with standard errors clustered at the county level in parentheses and school weights calculated as explained in Subsection E.3. The regressions are estimated on EIPL for the period from September 2020 to May 2021. Local affluence variables consist of zip-level household income, share of adults with College or higher education, share of dual-headed household with children. The school type fixed effects consists of indicators for charter school and non-charter school, and the school grade fixed effects consist of indicators for elementary vs. middle vs. high. vs. combined school for both samples. County controls consist of pre-pandemic ICU bed capacity, two-week lagged county COVID case and death rates, population density in the county, dummies for rural-urban continuum codes, maximum weekly temperature in the county, and dummies for the various non-pharmaceutical interventions.

where only the local affluence measure of interest is included in the regression along with the school share of non-white students and school type and grade controls. These regressions are thus similar to those reported in Table G1. Columns (2), (4), (6) and (8) of Table G6 add all the covariates included in our main regression, and are therefore comparable to results shown in the main regressions discussed in the text (Figure 4).

Table G6 confirms that an inverse relationship between EIPL and affluence holds with regards to incarceration rates and neighborhood poverty: public schools in areas with *higher* rates of incarceration and public schools in *poorer* neighborhoods (i.e. school with a higher index) provided on average *lower* EIPL during the pandemic. The relation with housing prices (as captured by average rents of two-bedroom apartments) is also consistent with out main results. The role of upward mobility is more difficult to fathom because there may not be a clear correlation between this indicator and the measures of local affluence considered in our main analysis. The coefficient is negative in the quasi-univariate regression, then turns positive when introducing the controls, but in absolute term the effect on EIPL is quite limited. Also, in Table G6 as in the main regression, the estimates of race become smaller after adding geographic controls, reflecting the fact that schools in suburban and town/rural areas provided on average higher EIPL, and suburban and town/rural areas are on average less affluent, have a smaller share of college-educated households, and have a smaller population of non-white students.

Table G6: Robustness check: Other indicators of local affluence

Dependent variable	Effective in-person learning (EIPL)							
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
OA upward mobility of 25th pctile children	-3.17***	0.92**						
	(0.61)	(0.40)						
OA incarceration rate of 25th pctile children			3.43***	0.78***				
			(0.47)	(0.24)				
OA rent for two-bedroom apartments					-8.35***	-1.64***		
					(0.52)	(0.50)		
EDGE school neighborhood poverty index							-6.36***	-0.38
							(0.34)	(0.30)
School share of non-white students	-23.22***	-5.79***	-22.92***	-5.99***	-18.26***	-5.57***	-23.95***	-5.99***
	(1.40)	(0.93)	(1.26)	(0.92)	(1.02)	(0.88)	(1.27)	(0.95)
Local affluence		✔		✔		✔		✔
School type and grade controls	✔	✔	✔	✔	✔	✔	✔	✔
Other school/district variables		✔		✔		✔		✔
County controls		✔		✔		✔		✔
R-squared	0.11	0.27	0.11	0.27	0.14	0.27	0.13	0.27
# of counties	2,863	2,633	2,863	2,633	2,863	2,633	2,863	2,633
# of schools	56,632	56,019	56,632	56,019	56,632	56,019	56,632	56,019
# of school districts	11,391	10,974	11,391	10,974	11,391	10,974	11,391	10,974

Notes: Each column reports coefficients from a weighted OLS regression on the public school sample, with standard errors clustered at the county level in parentheses and school weights calculated as explained in Subsection E.3. The regressions are estimated on EIPL for the period from September 2020 to May 2021. Local affluence variables consist of zip-level household income, share of adults with College or higher education, share of dual-headed household with children. The school type fixed effects consists of indicators for charter school and non-charter school, and the school grade fixed effects consist of indicators for elementary vs. middle vs. high. vs. combined school for both samples. Other school/district variables consist of mean math-RLA test scores, school size, school spending per student and ESSER funding per student. County controls consist of pre-pandemic ICU bed capacity, two-week lagged county COVID case and death rates, population density in the county, dummies for rural-urban continuum codes, maximum weekly temperature in the county, and dummies for various non-pharmaceutical interventions. OA (Opportunity Atlas) upward mobility is the mean household income rank for children whose parents were at the 25th percentile of the national income distribution where incomes for children are measured as mean earnings in 2014-2015 when they were between the ages 31-37; OA incarceration rate is the fraction of children born in 1978-1983 birth cohorts with parents at the 25th percentile of the national income distribution who were incarcerated in 2010.

G.2.7 Systematic regional differences. In order to understand better the role of the county-level regressors, Table G7, presents the results of introducing these regressors in isolation from each other. We begin in Column (1) with a regression on the teacher labor market, as measured by the unionization rate and cost index of hiring PK-12 educators, while controlling for pre-pandemic ICU bed capacity, two-week lagged county COVID case and death rates, dummies for various non-pharmaceutical interventions, maximum weekly temperature in the county, population density in the county, and county's rural/urban continuum codes. Column (2) focuses on the effects of political preferences. Note that the reason why the R-squared remains in the same ballpark is due to the county-level controls. Column (3) considers the effects of the COVID vaccination campaign together with the mask mandates. Last, Column (4) adds all five variables together. Foremost, Table G7 shows that the interrelation between the county-level regressors of interest is not so strong, perhaps with exception of the NPIs (the role of COVID vaccination rates is enhanced in Column (4) and that of mask mandates is reduced by half) and the local index of costs of hiring PK-12 educators (its role is dampened in Column (4)).

Table G7: Accounting for systematic geographical differences

Dependent variable	Effective in-person learning (EIPL)			
	(1)	**(2)**	**(3)**	**(4)**
Teacher unionization rate	-8.62***			-6.90***
	(1.05)			(1.01)
Local index of costs of hiring PK-12 educators	-9.58***			-3.93***
	(1.05)			(1.03)
Share of 2020 Republican voters		16.46***		13.45***
		(1.07)		(1.16)
Mask required in public			-14.40***	-7.85***
			(1.45)	(1.08)
COVID vaccination rate (two weeks lag)			4.93***	8.17***
			(0.63)	(0.56)
Local affluence	✔	✔	✔	✔
School type and grade controls	✔	✔	✔	✔
Other school/district variables	✔	✔	✔	✔
County health, pop. characteristics, weather, NPIs	✔	✔	✔	✔
R-squared	0.24	0.25	0.24	0.27
# of counties	2,633	2,633	2,633	2,633
# of schools	56,019	56,019	56,019	56,019
# of school districts	10,974	10,974	10,974	10,974

Notes: Each column reports coefficients from a weighted OLS regression on the public school sample, with standard errors clustered at the county level in parentheses and school weights calculated as explained in Subsection E.3. The regressions are estimated on EIPL for the period from September 2020 to May 2021. Local affluence variables consist of zip-level household income, share of adults with College or higher education, share of dual-headed household with children. The school type fixed effects consists of indicators for charter school and non-charter school, and the school grade fixed effects consist of indicators for elementary vs. middle vs. high. vs. combined school for both samples. Other school/district variables consist of mean math-RLA test scores, school size, school spending per student and ESSER funding per student. County health, pop.characteristics, weather, NPIs consist of pre-pandemic ICU bed capacity, two-week lagged county COVID case and death rates, population density in the county, dummies for rural-urban continuum codes, maximum weekly temperature in the county, and dummies for the various non-pharmaceutical interventions.

In results not reported here, we find that: (i) using the Republican vote share in the 2016 presidential election among our proxies for the general stance towards reopening schools barely change the results, which is unsurprising given the strong persistence in county-level Republican vote shares in the 2016 and 2020 presidential elections; (ii) changing the number of time lags used to measure COVID vaccination, infection and death rates matters for the coefficient on vaccination rates, while suggesting that a 2-weeks lag is appropriate;[12] restricting the sample to counties with at least 10 public schools, which reduces the sample size almost threefold, leave the results mainly unchanged.

[12]When using contemporaneous values of COVID vaccination, infection and death rates, the effect of vaccination rates is less pronounced – it is reduced by half –, and infection rates exert a negative effect on EIPL. On the other hand, with a lag of one month, the effect of COVID vaccination rates is very close to the baseline estimates. It is unclear how best to measure the dynamic relationships between the COVID health variables and EIPL, but in all instance the regressions show that the vaccination campaign is positively related to EIPL in a statistically and economically significant way.