Chaire en macroéconomie et prévisions

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

André Kurmann[1] et Etienne Lalé[2]

Cahier de recherche
Working paper
2022-03

Juillet / July 2022

ESG UQÀM
École des sciences de la gestion
Université du Québec à Montréal

[1] Drexel University
[2] Université du Québec à Montréal

Chaire en macroéconomie et prévisions

La Chaire en macroéconomie et prévisions est fière de s'appuyer sur un partenariat avec les organisations suivantes:

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

André Kurmann
Drexel University

Etienne Lalé
Université du Québec
à Montréal

July 23, 2022

Abstract

Social scientists have developed several schooling mode trackers to measure in-person, hybrid, and remote learning of students during the COVID-19 pandemic. In this paper, we compare eight of the most popular trackers for the U.S. and uncover substantial temporal and geographical differences, due in large part to how the trackers define the three schooling modes. We then estimate a new measure of effective in-person learning (EIPL) that combines information on school learning mode with cell phone data on school visits. The new measure provides a single number of the fraction of time that students spent learning in person and is made publicly available for a large, representative sample of both public and private schools. Consistent with other studies, we find that a school's share of non-white students and a school's prepandemic grades and size is associated with less in-person learning during the 2020-21 school year. Notably, we also find that schools in more affluent localities with higher pre-pandemic spending and schools receiving more federal emergency funding provided lower EIPL. These results are explained in large part by regional differences, reflecting political preferences, vaccination rates, teacher unionization rates, and local labor conditions.

JEL Classification: E24, I24

Keywords: COVID-19; School closures and reopenings; Effective in-person learning; Inequality

*Contact: Kurmann: Drexel University, LeBow College of Business, School of Economics, 3220 Market Street, Philadelphia, PA 19104 (email: kurmann.andre@drexel.edu); Lalé: Université du Québec à Montréal, Department of Economics, C.P. 8888, Succ. centre ville, Montréal (QC), H3C 3P8 Canada (email: lale.etienne@uqam.ca). Aseni Ariyaratne provided excellent research assistance. We thank Dennis Roche from Burbio, Nat Malkus from Return to Learn, and Safegraph for generously sharing their data. We also thank seminar participants at the 2nd joint IZA & Jacobs Center Workshop on "Consequences of COVID-19 for child and youth development" for their comments. The data constructed in this paper is available through the online repository of the Center for Open Science: http://doi.org/10.17605/osf.io/cghs2. All errors are our own.

1 Introduction

The COVID-19 pandemic led many schools in the U.S. to suspend or substantially reduce in-person learning. While available studies report conflicting results on the extent to which school closures helped prevent the spread of the virus,[1] evidence is emerging that remote instruction led to substantial learning losses and social-emotional harm with possibly large adverse long-term effects, especially for students from disadvantaged backgrounds.[2] Faced with the challenge of analyzing these consequences, different organizations have developed trackers to measure the amount of in-person, hybrid, and remote schooling that students obtained during the pandemic.

The objective of this paper is threefold. First, we compare existing schooling mode trackers and assess the extent to which they provide a consistent picture of in-person education during the pandemic. Second, we combine information from existing trackers with cell phone data on school visits to estimate a new measure of effective in-person learning (EIPL) for a large, representative sample of both public and private schools. Third, we study the association of EIPL with a host of school- and region-specific demographic and socio-economic indicators.

We compare eight of the most prominent trackers that collected information on schooling mode from school districts and state educational agencies or by directly surveying schools. We find large variations across the trackers in the percentages spent in the different schooling modes by region and time period. These variations, which are due to how each tracker defines the three schooling modes as well as coverage issues, also manifest themselves in differences in statistical associations of the trackers with regional, demographic, and socio-economic indicators. The variations pose an important challenge for researchers interested in documenting the extent and consequences of school closures across the U.S.

Motivated by these findings, we propose a new measure of the time that students effectively spent learning in person – EIPL. The measure maps anonymized cell phone data from Safegraph on visits to individual schools to information from schooling mode trackers. The mapping allows us to use school visits to estimate the fraction of hybrid learning that took place on an in-person basis and then select for each region the tracker information that provides the best fit with the visits data. The result is a single number of the relative importance of EIPL for the student population enrolled at a school. We estimate this number for a sample of more than 70,000 public and private schools that is highly representative of the universe of U.S. schools.

Compared to the existing schooling mode trackers, the EIPL measure has several advantages. It naturally addresses coverage issues with the different trackers; it is available for a large sample of both public and private schools at a weekly frequency for the entire 2020-21 school year; and the single number characteristic makes it more amenable to analyze the relationship with other variables than categorical indicators.[3] We make the EIPL dataset publicly available through the online repository of the Center for Open Science and hope it will be useful for other researchers.

[1] See Bravata et al. [2021], Chernozhukov et al. [2021a,b], Ertem et al. [2021], and the references therein, for systematic assessments of the effect of school closures on subsequent COVID-19 infections.

[2] See Dorn et al. [2021a], Halloran et al. [2021], Kogan and Lavertu [2021], Lewis et al. [2021], Goldhaber et al. [2022b] as well as Agostinelli et al. [2022], Jang and Yum [2020] or Fuchs-Schündeln et al. [2022] among others.

[3] As an example, consider two schools that experience the same increase in a tracker's in-person schooling mode but where one school was previously in remote mode and the other was previously in hybrid mode. The tracker's in-person measure would erroneously attribute the same increase in in-person learning to both schools, whereas our EIPL measure takes into account that hybrid schooling provides a certain amount of in-person learning. For further discussion and examples, see the main text.

In the last part, we investigate the extent to which pre-pandemic school and local characteristics correlate with EIPL during the 2020-21 school year. Naturally, these correlations should not be interpreted as causal, but they provide us with a set of stylized facts to understand the factors behind school closings and which segments of the student population were most affected. We find the following main results:

1. Public schools provided substantially less EIPL than private schools, with public charter schools ranking below public non-charter schools and private religious schools ranking above private nonreligious schools.

2. For both public and private schools, EIPL was lower in more affluent and more educated localities with a larger share of dual-headed households, and for schools with a larger share of non-white students.

3. For public schools, EIPL is negatively related to pre-pandemic school test scores, school size, and school spending per student as well as district Elementary and Secondary School Emergency Relief (ESSER) funding per student.

Finally, we show that a large part of these associations can be explained by a school's county share of Republican votes in the 2020 presidential election. COVID vaccination rates also predict higher EIPL while a state's teacher unionization rate and local labor conditions for teachers predict lower EIPL. In contrast, COVID case and death rates and whether a school is located in a dense, urban area do not have significant predictive power.

The relation of EIPL with race, school quality, size, spending per student, and Republican voting preferences validate results already documented in the literature (see, e.g., Gollwitzer et al. [2020]; Hartney and Finger [2020]; Parolin and Lee [2021]; Landivar et al. [2022]). This confirms that the proposed EIPL measure captures a component that is common across different school trackers. At the same time, our analysis establishes two results that, to the best of our knowledge, have not been highlighted before. First, we document a new nexus between income and public school closures: EIPL was on average lower – not higher – in more affluent localities; and this correlation is in large part accounted for by Republican voting preferences. Second, EIPL is negatively related to district ESSER funding per student, despite the fact that the program was advertised in Congress primarily as support for schools to reopen to in-person learning.

Taken together, the results suggest that at the national level, less affluent communities provided on average more EIPL. At the local level, however, schools that are on average associated with worse student outcomes provided less EIPL. These findings raise critical questions about education policy during the pandemic and the impact of in-person learning loss on future educational attainment as well as income inequality.

The paper is part of a growing literature that attempts to measure the extent and consequences of school closures during the pandemic. As reviewed below, several organizations and research teams have developed schooling mode trackers, among them Burbio, the Evidence Project, the COVID-19 School Data Hub, the Elementary School Operating Status team, and the American Enterprise Institute. In parallel, a number of studies have used school visits from anonymized cell phone data, in particular Safegraph, to proxy school closures during the pandemic: see Bravata et al. [2021], Chernozhukov et al. [2021a], Parolin and Lee [2021], Garcia and Cowan [2022], Hansen et al. [2022]. Our methodological contribution consists in combining these two data sources to develop a single number of the time effectively spent learning in person that we estimate for a large sample of schools. As described above, this number has clear merits over the existing trackers, not just because of its granularity and coverage, but because categorical variables can provide a misleading measure of in-person learning as reflected by the large variations across trackers.[4] In turn, the use of cell-phone based school visits data alone is subject to issues of interpretation and measurement error. First, without additional information, it is not clear what a given decline in visits to a school during the pandemic represents in terms of lost in-person learning.[5] Second, attributing cell phones to a particular location is challenging, and our analysis reveals that this leads to very noisy data for a non-negligible number of schools. When building our EIPL measure, we therefore impose stringent quality checks and only retain schools for which the visits data is reliable.

Our EIPL measure can be used as an input for numerous analyses of the effects of pandemic school closures, such as on COVID infections and deaths (Auger et al. [2020]; Bravata et al. [2021]; Chernozhukov et al. [2021a]; Ertem et al. [2021]; Goldhaber et al. [2022a]); on local labor market (Garcia and Cowan [2022]; Hansen et al. [2022]; Landivar et al. [2022]; Prados et al. [2021]); on public school student enrollment (Dee et al. [2021]); on student learning achievement (Dorn et al. [2021b]; Halloran et al. [2021]; Engzell et al. [2021]; Kogan and Lavertu [2021]; Goldhaber et al. [2022b]), and on long-term income inequality and welfare (Agostinelli et al. [2022]; Jang and Yum [2020]; Fuchs-Schündeln et al. [2022]).[6] This literature is still at an early stage and rapidly expanding. In this respect, our regression results of EIPL against a rich set of local and regional indicators may serve as a useful guide for further analysis.

The paper proceeds as follows. Section 2 is our review and comparison of information provided by the schooling mode trackers. Section 3 explains our empirical approach for measuring EIPL based on learning mode data combined with school visits. The correlation study of EIPL with local population, school, and regional indicators is presented in Section 4. Section 5 concludes.

2 Comparison of schooling mode trackers

This section provides a comparison of what we consider to be some of the most prominent schooling mode trackers. To our knowledge, there exists no systematic review to help guide researchers interested in using these data to study pandemic school closures.

We consider schooling mode trackers that were constructed from guidelines and/or practices posted publicly by school districts and state educational agencies or by directly surveying schools. We do not impose any restriction on the mode of data collection, which may differ in terms of frequency, systematicity, and sampling methods. However, we do require that the primary data collected is from direct source of information about school learning modes. We require that the data are at a sub-national level.

We use several tools to conduct a wide search of available schooling mode trackers. We use Scopus's document search tools for the following keywords: "COVID-19", "School closure", "School reopening", and

[4] A blogpost by Camp and Zamarro [2022] also discusses issues of comparability for three of the schooling mode trackers reviewed below. Our work substantially expands over this discussion.

[5] Suppose cell phone usage is concentrated among school staff and parents. If school staff returned to school more quickly than students (e.g. to prepare the return of students or to teach only some students in person while others remained in remote-learning mode) or if students get dropped off and picked up by parents instead of using buses, then foot traffic data alone would overestimate in-person learning during the pandemic. Vice versa, consider a school that contains a playground or sports fields that are usually open to the public or used for games. If due to the pandemic, access to this playground or sports field is restricted even though the school has reopened, then cell phone traffic alone would underestimate in-person learning.

[6] Fuchs-Schündeln et al. [2022] represents a first application, using an earlier version of our EIPL. The version here is substantially expanded, exploiting additional information from schooling mode trackers across different regions and time periods.

"Schooling mode" (filtering on the date when they first appeared). We conduct a similar search on Mendeley, the Center for Open Science, and the open ICPSR data repository. We manually check the first 100 papers matching those keywords on Google scholar. Finally, we search on Google as well as on several education blogs to find data that may not have been used in academic research.

We end up with eight trackers that fit our definition: Burbio, the Center on Reinventing Public Education (CRPE), the COVID-19 School Data Hub (CSDH), Education Week (EdWeek), 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).

Table 1 overviews the trackers in terms of data structure and coverage, data source and collection methods, and available measures. There are a number of key differences across trackers. 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. In particular, depending on the tracker, hybrid learning may refer to part-day, part-week, rotating weeks, or alternating grades; and the fraction of part-day / part-week to qualify for hybrid learning varies (see Table 1 of the online appendix for details).

Table 1: 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. • CSDH team communicated with the state for clarification when questions arose regarding data inconsistencies, missing information, etc.	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.

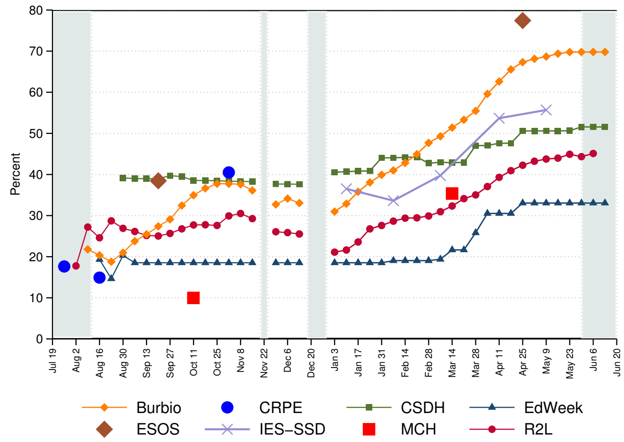
Figure 1 compares the schooling modes of 2020-21 according to the different trackers.[7] While they generally evolve in the same direction, there are also large differences in both magnitude and timing. In particular, there is substantial disagreement about the extent to which schools returned to in-person learning, and whether this return was mainly due to a shift away from hybrid or remote learning. Moreover, while Burbio, R2L, and IES-SSD predict a relatively smooth transition out of remote learning during the Spring term, the CSDH and EdWeek report more abrupt changes, likely due to the frequency of the data collection process. In sum, the eight trackers paint a qualitatively similar but quantitatively different picture of school closures and the provision of in-person learning during the pandemic.

Aside from the large variations in average schooling mode over time, there are also large geographic

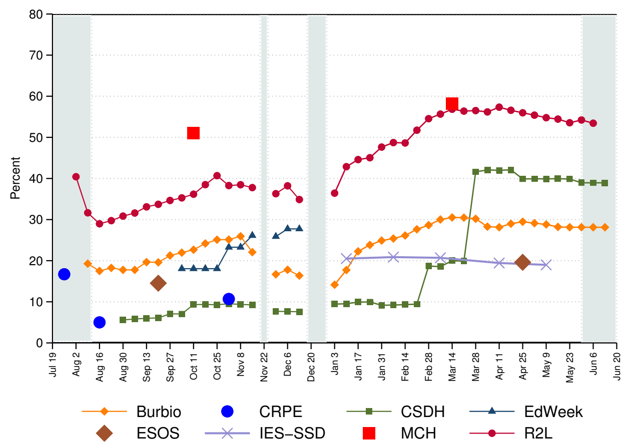
[7]Figure 1 presents the student-weighted average of each learning mode from each tracker. For CSDH, we average data using school-level and district-level student enrollment data. For schools districts that also have school-level data available in CSDH, we first aggregate data to the district level by 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.

Figure 1: Comparison of Schooling Mode Trackers

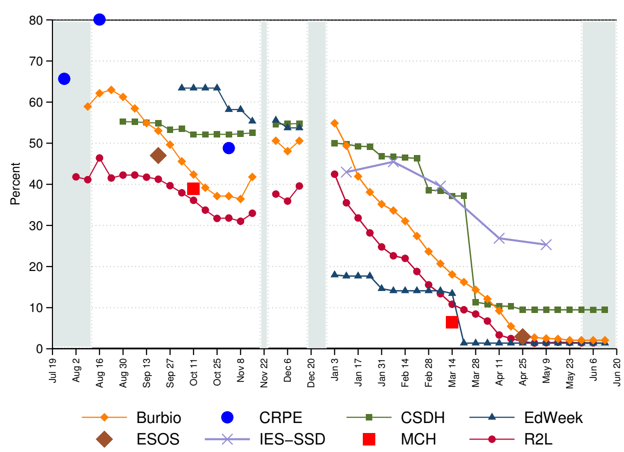
(a) In-person learning



(b) Hybrid learning



(c) Remote learning



Notes: The figures show the percentage share of each learning mode according to the different trackers, aggregated using public school student enrollment at either the school, district, county, or state level. The learning mode trackers are: Burbio, the Center on reinventing public education (CRPE), the COVID-19 school data hub (CSDH), EducationWeek (EdWeek), 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). The shaded regions denote the Summer, Winter, and Thanksgiving breaks.

variations between the trackers. Consider for instance Burbio and R2L, which according to 1 appear to be fairly well correlated with each other. Yet, across all counties and weeks of the 2020-21 school year, the correlation between Burbio's and R2L's share of in-person learning is 0.59, and the correlation for hybrid learning is 0.45. Digging deeper, the county-level correlation over time of the two trackers' share of in-person learning is *negative* for about 20% of all counties, with an interquartile range of correlations of 0.78 (similar results obtain for the share of hybrid learning). For the other trackers, the regional disparities in schooling mode are even larger.

While a detailed analysis is beyond the scope of this paper, our assessment of the data suggests the definition of the three schooling modes and in particular hybrid learning is the main reason for the differences between school trackers. In addition, the trackers vary substantially in terms of coverage across regions, which leads to both geographic and average variations over time. These differences pose an important challenge for the study of school closures across the U.S. and how the resulting loss of in-person learning affects school enrollment and learning outcomes of students from different backgrounds. Indeed, as we explain below, the variations in schooling mode trackers lead to large differences in the statistical association of each tracker with local school and socio-economic indicators as well as regional characteristics.

3 From changes in school visits to effective in-person learning

This section describes how we use school visits data together with information from different schooling mode trackers to construct our EIPL measure. Details and additional analysis are provided in the online appendix.

3.1 Data, sample restrictions, and measurement of changes in school visits

A key input for our EIPL measure comes from Safegraph, a private company which provides data on over 7 million Places of Interest (POIs) for the U.S., including visits derived from over 40 million anonymized cell phones. We retain all POIs with North American Industry Classification System code 611110 ("Elementary and Secondary Schools") that have weekly visit data. We match these POIs by school name and geo-location or address to the universe of public and private schools from the NCES's Common Core of Data and the Private School Universe Survey, which results in about 110,000 high-quality matches. Relative to the universe of schools, we lose about 12,000 schools. The matched sample remains highly representative in terms of demographic and geographic makeup.

The weekly visit count for each school is organized in seven dwell time intervals, ranging from less than 5 minutes to more than 240 minutes. Visits decline during major holidays and summer break; drop precipitously on March 13, 2020 and remain substantially lower on average thereafter. At the same time, due to the increase in cell phones covered by Safegraph, visits generally trend upward prior to the pandemic; and visits to individual schools can be subject to substantial variation, both from one week to another and across dwell time intervals. While some of these variations reflect school characteristics and idiosyncratic events, others are due to the inherent difficulty of attributing cell phones to a particular location.

To address these measurement issues, we proceed in three steps. First, we construct a dwell-time weighted average of weekly visits for each school that is normalized by the weekly count of cell phones covered by

Safegraph at the state level:

$$\widetilde{v}_{j,t} = \frac{1}{n_{s(j),t}} \sum_{d=1}^{7} \omega_j(d)\, v_{j,t}(d)\,, \tag{1}$$

where $v_{j,t}(d)$ denotes raw visits of dwell time d for school j in week t; $\omega_j(d) = \frac{\sum_{t=t_{-1}}^{t_0} v_{j,t}(d)}{\sum_{t=t_{-1}}^{t_0} v_{j,t}}$ measures the importance of visits of dwell time d for school j during reference period $t = t_{-1}, \ldots, t_0$ beginning in November 2019 through the end of February 2020 (excluding the weeks of Thanksgiving, Christmas and New Year); and $n_{s(j),t}$ is the normalization by the count of devices in state $s(j)$ in which school j is located.

Second, we drop about 37,000 schools with sparse or noisy visit data – an issue that seems to be overlooked by existing analyses of the Safegraph data. This step is important in terms of sample construction. For the approximately 73,000 schools that remain, we construct sample weights that keep the dataset representative of the universe of schools. We estimate these weights using a Probit model where the dependent variable is an indicator taking the value of 1 for schools that are included in the dataset and is 0 for schools not matched to Safegraph or with sparse or noisy visit data; and the regressors are county-level measures of family structure and education, population density, geographic locale of a school, and Census region dummies. The estimated school-level sample weights are included in the public release of our dataset.

In the final step, we construct the change in school visits as the percent difference in dwell-time weighted, normalized visits relative to the average $\widetilde{v}_{j,0} = \frac{1}{t_0 - t_{-1} + 1} \sum_{t=t_{-1}}^{t_0} \widetilde{v}_{j,t}$ over the reference period:

$$\Delta\widetilde{v}_{j,t} = 100 \times \frac{\widetilde{v}_{j,t} - \widetilde{v}_{j,0}}{\widetilde{v}_{j,0}}. \tag{2}$$

3.2 Estimating effective in-person learning

To estimate EIPL, we map changes in Safegraph school visits, $\Delta\widetilde{v}_{j,t}$, to changes in learning mode from Burbio and R2L. The reason we select the Burbio and R2L schooling mode data is that they provide high-frequency (i.e. weekly) variations that enable us to measure the substitution across learning modes and thereby extract the component of hybrid learning that is effectively done in person.[8] Our approach proceeds in two steps. First, we aggregate changes in weekly visits to the county level for Burbio, respectively the district level for R2L, and estimate the mapping. Second, we use the estimates to predict EIPL at the individual school level.

For expositional purposes, we focus on the county-level aggregation; the steps for the district-level aggregation are analogous. Denote by $\Delta\widetilde{v}_{c,t} = \sum_{j \in c} \kappa_j \Delta\widetilde{v}_{j,t}$ the average change in school visits in county c in week t relative to the reference period, where κ_j is the share of county c's students enrolled in school j. Next, define the fraction of total school time that is EIPL by students in county c during week t as

$$EIPL_{c,t} = T_{c,t} + \gamma H_{c,t}, \tag{3}$$

where $T_{c,t}$ is the share of school time in traditional in-person learning mode, $H_{c,t}$ is the share in hybrid learning mode, and γ defines the fraction of hybrid learning spent in person.

Since both $\Delta\widetilde{v}_{c,t}$ and $EIPL_{c,t}$ measure percent deviations from the pre-pandemic baseline, we can

[8] See Table 1. Weekly data are also available from CSDH and EdWeek. However, the weekly CSDH data covers about 10,000 schools (most schools in CSDH are observed at the monthly frequency), and the EdWeek data is at the state level.

formulate the relationship between the two variables as $EIPL_{c,t} = \alpha + \beta\Delta\widetilde{v}_{c,t} + \varepsilon_{c,t}$, or equivalently

$$T_{c,t} = \alpha + \beta\Delta\widetilde{v}_{c,t} - \gamma H_{c,t} + \varepsilon_{c,t}. \tag{4}$$

The regression tells us not only how a given change in school visits maps into EIPL, but also the proportion γ of hybrid learning spent in person.

While (4) can be estimated as a panel across all counties, respectively school districts, one important concern is that the mapping between changes in school visits and learning modes may not always be the same; e.g. because of differences in hybrid learning arrangements across districts. A second concern is that the quality of the Burbio and the R2L data differs across region and time period. To address these concerns, we estimate (4) separately for Burbio and R2L at both the Core-Based Statistical Areas (CBSA) level and the state level. For each regression, we restrict $\alpha = 100$ as implied by the pre-pandemic baseline when schools were fully in person (i.e. $T_{c,0} = 100$, $\Delta\widetilde{v}_{c,0} = 0$, and $H_{c,0} = 0$), and find the regression time window with the highest R-squared. This provides us with four sets (Burbio-CBSA, R2L-CBSA, Burbio-state, R2L-state) of estimates $\left\{\hat{\beta}, \hat{\gamma}\right\}$ and associated R-squared for every school j. To predict $EIPL_{j,t}$, we then use the regression coefficients associated with the highest R-squared subject to the restrictions that the proportion of hybrid learning spent in person is within its theoretical bounds, $0 < \hat{\gamma} < 1$, and the R-squared is over 0.25.[9] Using the coefficient $\hat{\beta}$ obtained in this way, we compute EIPL at school j during week t as: $\widehat{EIPL}_{j,t} = 100 + \hat{\beta}\Delta\widetilde{v}_{j,t}$. Observe that the algorithm effectively trades off geographic variation in regression coefficients with regression fit and does so by using either the Burbio or the R2L data that is of higher quality.[10]

Table 2: Mapping school visits to Effective In-Person Learning

(a) Source of regression coefficients to map school visits to EIPL

	Burbio (CBSA)	R2L (CBSA)	Burbio (State)	R2L (State)
Number of schools	21,615	17,098	16,166	18,348
Percent of schools	29.5	23.3	22.1	25.1

(b) Distribution of regression coefficients to map school visits to EIPL

	Mean	Percentile				
		5th	25th	50th	75th	95th
$\hat{\beta}$	1.18	0.72	1.13	1.20	1.27	1.43
$\hat{\gamma}$	0.29	0.04	0.17	0.27	0.37	0.63
R squared	0.81	0.45	0.68	0.88	0.96	0.99

Notes: Panel (a) shows the distribution of schools by type of regression coefficient retained for the OLS estimation of (4). Panel (b) shows the distribution of retained regression coefficients and R-squared, weighted by the different school weights.

Table 2 reports summary statistics for the estimated coefficients retained to predict $EIPL_{j,t}$. As shown

[9] The restrictions on $\hat{\gamma}$ applies only in a few cases. R-squared is lower than 0.25 for all four regressions in Arkansas and Maine, where both the Burbio and R2L data appear to be of low quality. For those, we use regression coefficients from neighboring states.

[10] We could apply the same algorithm at the county level. In many cases, however, this would result in R-squared that are lower than at the CBSA level. More generally, we have experimented with several modifications of the algorithm, and the results reported below remain very robust.

in panel (a), across the approximately 73,000 schools, the retained estimates are evenly distributed between Burbio and R2L and between the CBSA and state level, indicating that both Burbio and R2L data are useful and that allowing for finer geographic variation would in about half the cases produce a worse fit.[11]

Panel (b) shows the school-weighted distribution of the retained estimates and R-squared across schools. The regressions are generally tightly estimated with a median R-squared of 0.88. The estimated mapping between changes in school visits and EIPL, $\hat{\beta}$, ranges from about 0.7 to 1.4 while the estimated proportion of hybrid learning spent in person, $\hat{\gamma}$, ranges from 0.04 to 0.6. As confirmed in further analysis, these distributions reflect large regional variations in the mapping from school visits to EIPL. Focusing on means, a one percentage point decline in school visits predicts an average reduction in EIPL by 1.2 percentage points, and the predicted fraction of hybrid learning spent in person is 0.3 or the equivalent of 1.5 days out of a 5 day school week.

3.3 Relation of schooling mode trackers with EIPL

It is instructive to compare EIPL with the schooling mode trackers. In Table 3, 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 3: 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, the correlations with EIPL are similar, suggesting that EIPL captures a common

[11] One may be concerned that the Safegraph data is more noisy in some parts of the country, but the pre-pandemic variance of visit changes is evenly distributed across CBSAs.

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 3 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/middle/high), 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.

4 EIPL during the pandemic: when, where, and for whom?

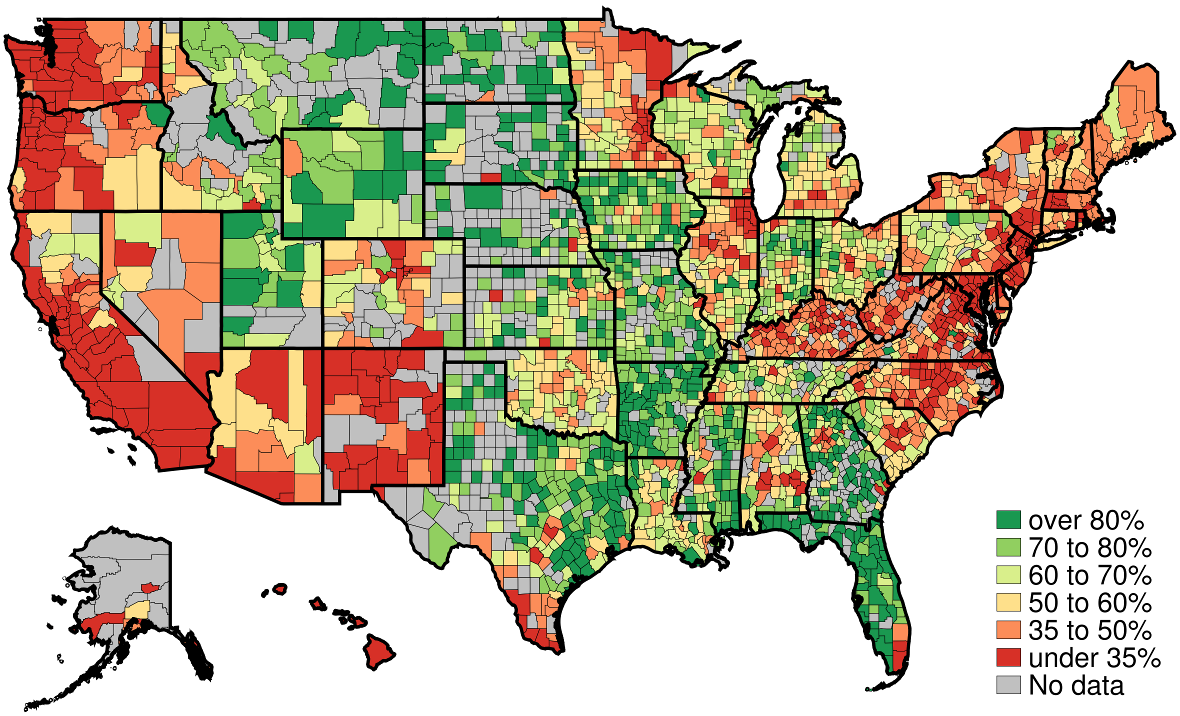
To illustrate the properties of our EIPL measure, we document differences in EIPL over time and across regions, and then turn to a host of characteristics that may account for those differences. Several results in this section have been established in prior studies, suggesting that the EIPL measure captures a component that is common across many schooling mode and schools' cell phone traffic databases. We clearly indicate those results and explain which of the findings are novel.

Panel (a) of Figure 2 describes the regional disparities in EIPL. After dropping to between 0% and 20% for almost all counties from March to May 2020 (see the appendix), EIPL recovered to 60% or higher in many counties in the South and Central North, while in counties in the North and Mid-Atlantic and the West it remained stuck in the 0% to 35% range. Panel (b) provides further information by reporting the top 10 and bottom 10 cities in terms of average EIPL from September 2020 through May 2021 among the 50 biggest U.S. cities by population. In cities in Florida such as Jacksonville, Tampa or Orlando, EIPL averaged over 75%, whereas in cities in California, Oregon and Washington such as Los Angeles, Portland or Seattle, EIPL averaged 20% or less.

Why did some schools return to in-person learning more quickly than others and what explains the large regional disparities? To answer this question, we analyze how various observable school and local characteristics surrounding the school correlate with EIPL. Then we return to geography and examine the extent to which these are driven by systematic regional differences. This correlational study highlights a set of facts that can help us understand the "for whom" in order to quantify the consequences of pandemic-induced learning losses for different segments of the student population and formulate appropriate policies going forward.

Figure 2: Regional Disparities in Effective In-Person Learning

(a) Effective in-person learning across U.S. counties during the 2020-21 school year



(b) The top 10 and bottom 10 U.S. cities in terms of effective in-person learning

Rank	CBSA name	EIPL	Rank	CBSA name	EIPL
1	Jacksonville, FL	87.6%	41	Sacramento-Arden-Arcade-Roseville, CA	22.9%
2	Tampa-St. Petersburg-Clearwater, FL	81.3%	42	Washington-Arlington-Alexandria DC-VA	22.9%
3	Orlando, FL	77.1%	43	Baltimore-Towson, MD	22.4%
4	Houston-Baytown-Sugar Land, TX	61.9%	44	Seattle-Bellevue-Everett, WA	20.3%
5	Fort Worth-Arlington, TX	61.8%	45	Portland-Vancouver-Beaverton, OR-WA	20.2%
6	Cincinnati-Middletown, OH-KY-IN	61.3%	46	San Jose-Sunnyvale-Santa Clara, CA	17.0%
7	Dallas-Plano-Irving, TX	60.7%	47	Las Vegas-Paradise, NV	16.5%
8	Detroit-Livonia-Dearborn, MI	58.3%	48	Los Angeles-Long Beach-Santa Ana, CA	16.4%
9	Nassau-Suffolk, NY	57.6%	49	Oakland-Fremont-Hayward, CA	15.9%
10	Nashville-Davidson--Murfreesboro, TN	57.4%	50	Riverside-San Bernardino-Ontario, CA	14.5%

Notes: The top panel shows the student-weighted average county EIPL from September 2020 to May 2021 by different percentile ranges for all counties for which we have reliable data on at least three schools. The bottom panel shows the top-10 and bottom-10 Core-Based Statistical Areas (CBSAs) in terms of average EIPL among the 50 largest CBSAs by population. EIPL for each CBSA is computed as the student-weighted average across schools with reliable data.

4.1 School type and grade

Perhaps the most obvious observable school characteristics are type (public non-charter, public charter, private non-religious, or private religious school) and grade (elementary school, middle school, high school, or a combination thereof) as designated by the NCES. From March to May 2020, all school type and grade combinations averaged less than 20% of EIPL. For the 2020-21 school year, however, there are substantial differences. EIPL is lowest for public charter schools (averaging 36%), followed by public non-charter schools (44%), private non-religious schools (51%), and private religious schools (57%). In turn, EIPL is lower for middle and high-schools (averaging 39%) than for elementary schools (56%), and these differences are more pronounced for public than for private schools.

The EIPL ranking by school type, which to our knowledge has not been analyzed in prior research, may come as a surprise for two reasons. First, public charter schools are typically independent and not unionized whereas public non-charter schools belong to school districts that, for some urban areas, are comprised of several hundred schools and often unionized. One could have expected that these features would have made it easier for charter schools to reopen to in-person learning. Second, according to Hanson [2021], tuition for non-religious private schools is on average more than twice as high as tuition for religious private schools. The additional resources and resulting smaller class sizes could have made it easier for non-religious private schools to reopen to in-person learning. Yet, in both cases, exactly the opposite occurred.

The EIPL ranking by school grade is not a new finding; see Parolin and Lee [2021], Musaddiq et al. [2021], or Burbio's dashboard. Given the importance of the early stages of schooling for human capital accumulation, it likely contributed to shielding younger children from some of the adverse effects of school closures, compared to their older peers. Indeed, Fuchs-Schündeln et al. [2022] structurally estimate that it is children just starting secondary school during the 2019-20 school year that endure the largest losses in their earnings capacity in the long run.

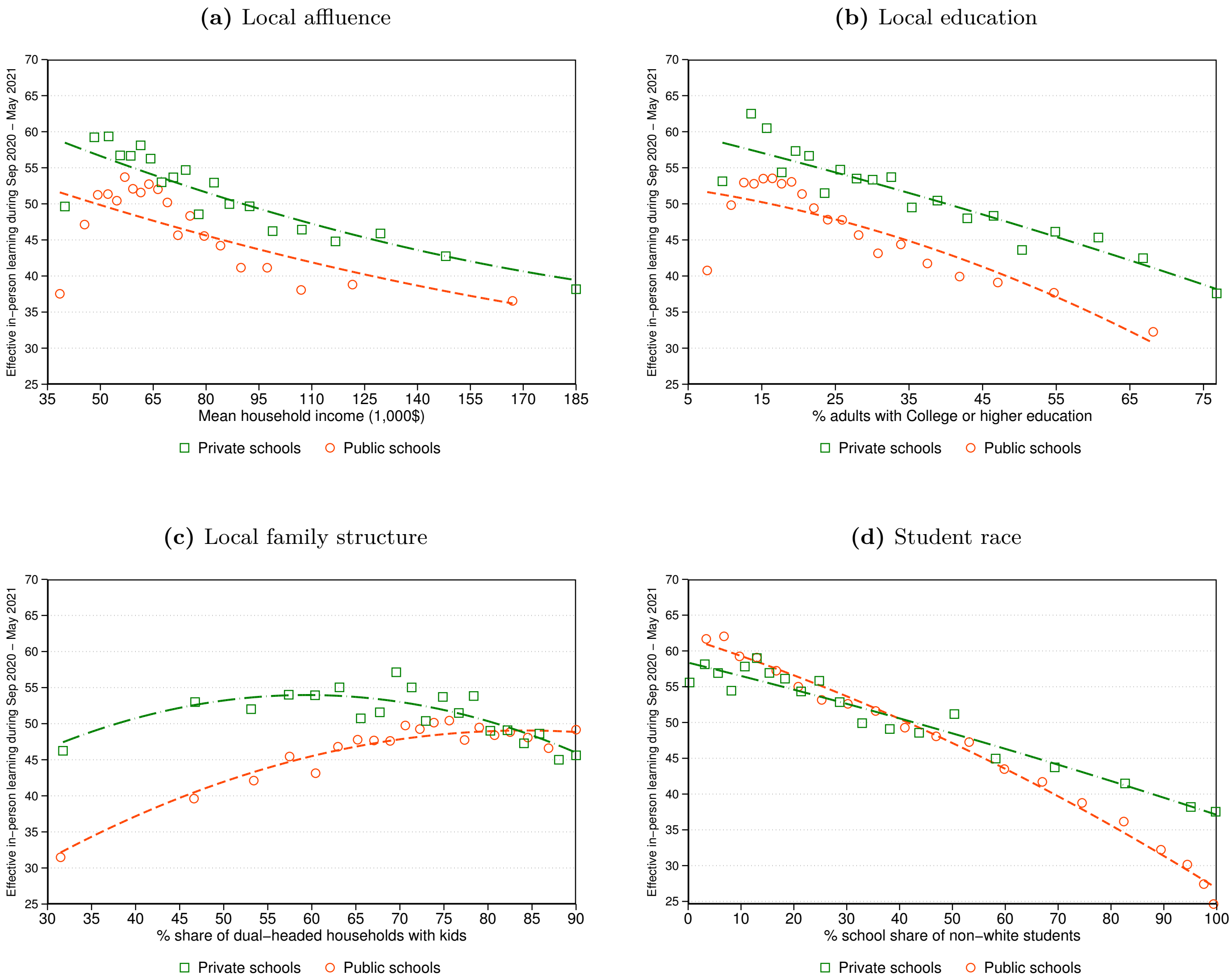
4.2 Local affluence and education, family structure, and student race

Next we consider EIPL by local affluence, education and family structure, which are prominent inputs for models of human capital accumulation (e.g. Cunha et al., 2010), as well as student race. We proxy local affluence by household income, education by the share of households with a college degree or higher, and family structure by the share dual-headed households with children, all measured at the zip-code level of the school and based on 2016-2019 estimates from the American Community Survey. For race, we use the school's share of non-white students as provided by the NCES. Results are robust to using the variables at the census block group or tract of the school, or at the school district level.[12]

Figure 3 reports binned scatterplots of the unconditional relationship between average school EIPL from September 2020 to May 2021 and the different variables. As shown in panel (a), there is an inverse relationship between EIPL and household income: schools in zip-codes with *high* household income provided on average *lower* EIPL. Panel (b) shows a similar result for education: schools in zip-codes with a *high* share of college-educated people provided on average *lower* EIPL. Consistent with the above results, EIPL is on average about 10% higher for private schools than for public schools but the relationship of EIPL with

[12] We have also considered several alternative socio-economic indicators describing the neighborhood of the school, including Chetty et al. [2020]'s indicators of Income Mobility and many of the other variables available from their Opportunity Atlas database. While these variables are correlated with EIPL, they are also highly correlated with the above measures of local affluence and education and do not add significant explanatory power to the below regressions.

Figure 3: Effective In-Person Learning by Local Affluence, Education, Family Structure, and Race



Notes: The figures show binned scatterplots of average EIPL from September 2020 to May 2021 for public schools and private schools, respectively, by (a) zip-code average household income, (b) zip-code average share of household with a college degree or higher, (c) zip-code share of dual-headed households, and (d) school share of non-white students. Observations are weighted by the school-specific sampling weight described in the appendix.

household income and education is otherwise very similar. Panel (c) shows that there is positive relationship between EIPL and local share of dual-headed households for public schools but no systematic relationship for private schools. As we will see below, the positive relationship for public schools changes once we control for other observable school characteristics. Finally, as shown in panel (d), there is also a clear inverse relationship between EIPL and the share of non-white students. For schools with close to 0% of non-white students, EIPL averaged over 60%, independent of whether the school is public or private. For schools with close to 100% of non-white students, in contrast, EIPL averaged only about 25% for public schools and just below 40% for private schools.

Given the general association of poverty with race, the inverse relationship of EIPL with both affluence and race may come as a surprise.[13] We further investigate this result through OLS regressions of average 2020-21 school EIPL on the different measures. Figure 4 reports the results. To save on space, we show only results for public schools, but the results for private schools are very similar. The brown square-shaped plots show the point estimates and 95% confidence intervals from regressing EIPL separately on zip-level household income, education, and share of dual-headed households together with the share of non-white students and controls for school type and school grade. The yellow and red plots are discussed further below. The coefficients are scaled so that they show the implied change in EIPL of going from the 25th to the 75th percentile of the distribution of a variable.[14]

All three indicators of affluence are negatively related to EIPL in a significant and quantitatively important manner. EIPL for a school located in a zip-code at the 75th percentile of the income and education distribution was on average 5%, respectively 7-8% lower than for a school at the 25th percentile. The relationship between and the share of dual-headed households with children is also negative, due to the fact that schools in zip codes with a larger share dual-headed households have on average a smaller share of non-white students. This highlights the importance of analyzing the relation between EIPL and different school characteristics in a multivariate setting. To our knowledge, these results have not been highlighted by other studies. They indicate that at the national level less affluent communities provided on average more – not less – EIPL.

The negative relationship between EIPL and a school's share of non-white students, on the other hand, is not a new finding. Even after conditioning on local income, education, and parental structure, EIPL for a school with a student body at the 75th percentile of the non-white distribution was on average 15-22% lower during the 2020-21 school year than for a school at the 25th percentile of the distribution. Further analysis shows that this negative relationship is in large part driven by the share of Hispanic students and less by the share of black students. These findings are consistent with Camp and Zamarro [2022], Landivar et al. [2022], and Parolin and Lee [2021], among others.

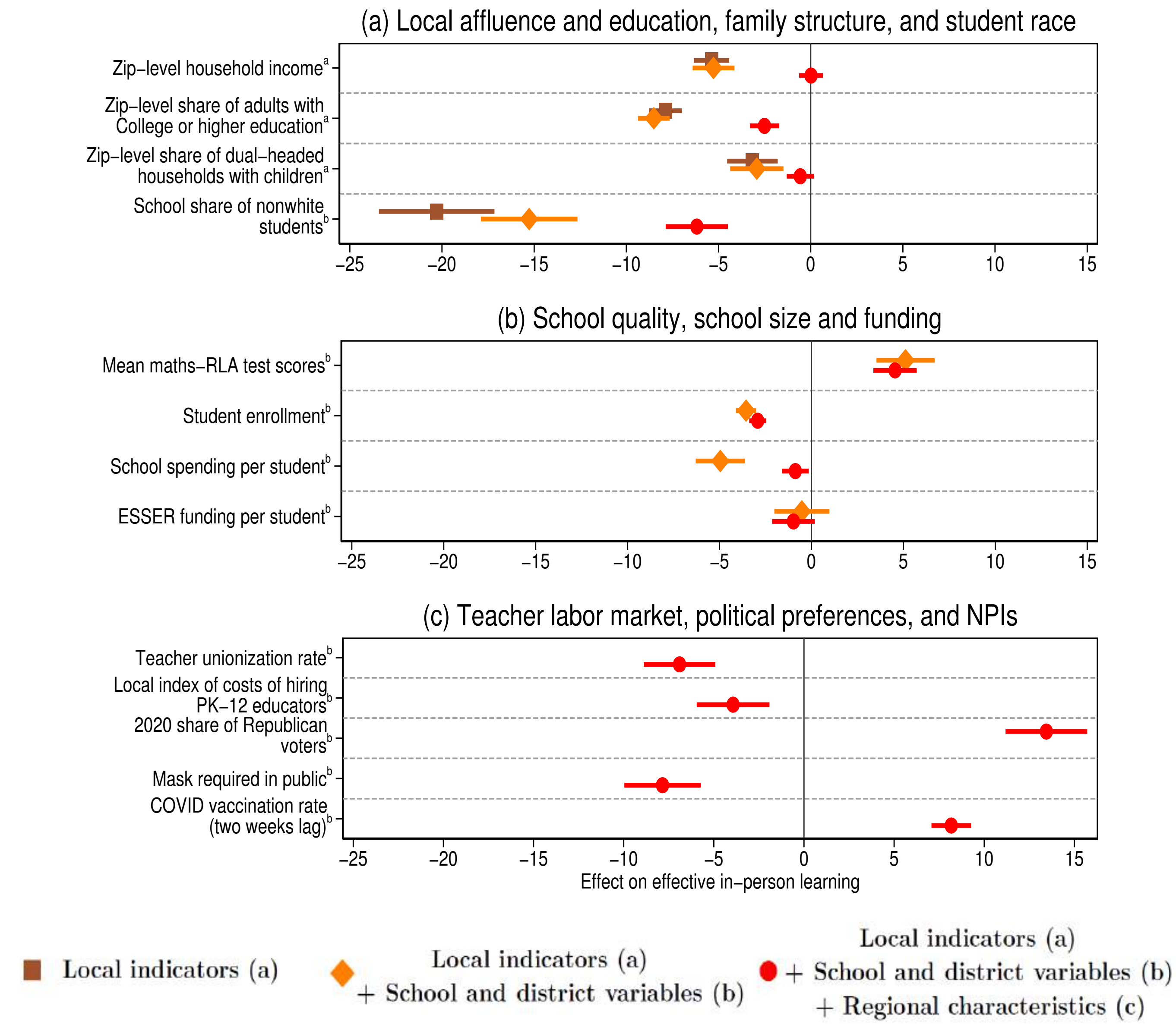
4.3 Public school tests scores, school size, and school funding

We extend the analysis by considering the relation between EIPL and school quality, school size, and school funding. We proxy school quality with pre-pandemic average tests scores, based on data from Fahle et al.

[13] As shown in the appendix, the share of a school's non-white students is essentially uncorrelated with local affluence, education, and parental structure.

[14] More precisely, all right-hand side variables are expressed as deviations from the mean, normalized by the interquartile range. Here and below, all estimates are weighted by the school-specific sampling weights to ensure representativeness; standard errors are clustered at the county level.

Figure 4: The Relationship of Effective In-Person Learning with School and Local Characteristics



(*) All regressions include School type and grade controls

Notes: The figure shows the estimated effects on EIPL from weighted OLS regressions with standard errors clustered at the county level and school weights calculated as explained in the appendix. The sample consists of approximately 60,000 public schools. The regressions are estimated for weekly school EIPL from September 2020 to May 2021. The estimates for the first three variables, denoted by ^{a}, 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 denoted by ^{b} are the result of regressions where all the variables are included jointly. The brown square-shaped estimates show the effects of regressing EIPL on the variables in the top box only. The yellow diamond-shaped estimates show the effects of regressing EIPL on the variables in the top and middle box. The red round-shaped estimates show the effects of regressing EIPL on the variables in all three boxes. All regressions control for school type (charter vs. non-charter school) and school grade (elementary vs. middle vs. high. vs. combined school). In addition, the regressions for the red round-shaped estimates 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 so that they show the implied change in EIPL of going from the 25th percentile to the 75th percentile of the distribution of a variable. See the appendix for details.

[2021].[15] For school funding, we consider both pre-pandemic school spending per student obtained from EdunomicsLab [2021] as well as district-level ESSER funding by student compiled by Malkus [2021b]

As above, we estimate OLS regressions of EIPL on local income, education, parental structure, and student race and then add the different variables.[16] As shown by the yellow diamond-shaped coefficient plots in Figure 4, EIPL is estimated to be higher for schools in districts with higher pre-pandemic test scores. Parolin and Lee [2021] obtain the same result based on pre-COVID third grade math test scores from Chetty et al. [2014]'s Opportunity Atlas database. EIPL is lower for larger schools, and interestingly, is also inversely related to school spending and ESSER funding per student, although not significantly so in the latter case. Note that these estimates control for local income, education, parental structure, and race, suggesting that school quality, school size, and school funding are independent predictors of EIPL.

The negative association between EIPL and school spending is a new finding, as far as we know, but it may have been expected given the positive correlation between school spending and local affluence. The absence of a positive association of EIPL with ESSER funding per student, in turn, is remarkable because ESSER, which was appropriated by Congress in three waves totaling $190 billion or almost five times the annual federal K-12 spending prior to the pandemic, was advertised primarily as support for school reopening. We return to this point below. Also note that controlling for test scores, school size, and school funding barely changes the negative relation of EIPL with local affluence, education, and parental structure, but reduces the effect of share of non-white students by roughly one third. This is because the share of non-white students is negatively correlated with test scores and positively correlated with school size. Even so, the share of non-white students remains a strong predictor of lower EIPL.

4.4 Geography

For the last part of the analysis, we ask how much of the relation of EIPL with school and local characteristics is driven by systematic regional differences that are not directly related to the school. From Figure 2 we know that schools with higher EIPL are generally located in the central and southern parts of the U.S. – regions that in general were more favorable towards reopening the economy despite potential health risks and at the same time have seen lower COVID vaccination rates. To assess the relative importance of these factors for EIPL, we re-estimate the above regressions and add county-level vote shares in the 2020 presidential election as a proxy for the general stance towards reopening the economy and two-week lagged vaccination rates. Furthermore, we add teacher unionization rates, a comparable wage index for PK-12 educators in the local labor market, controls for a county's COVID health situation as measured by pre-pandemic ICU bed capacity, lagged COVID case and death rates and various non-pharmaceutical interventions (NPIs), as well maximum weekly temperature and indicators of urban density.

The red round-shaped coefficient plots in Figure 4 report the results. First, schools in areas with higher teacher unionization rates and a higher comparable wage index provided on average lower EIPL. At the same time, a large share of Republican votes in the 2020 presidential election is a strong predictor of higher EIPL during the 2020-21 school year.[17] These results are not new. Hartney and Finger [2020] report similar

[15] We use district-level average test scores for 2018-19, which are available for almost all districts. School-level test scores are available for only a subset of schools and yield very similar results.

[16] As above, we include local income, education, and parental structure one-by-one together with race and the other variables. Since the estimates for race and the other variables barely change across regressions, we report these estimates controlling for local income, education, and parental structure jointly.

[17] Partisan vote shares are highly persistent at the county level. Hence, results are virtually identical when using vote shares

correlations between teachers' union strength and the presidential vote shares of the Republican candidate on the one hand and a district's propensity to reopen schools to in-person learning on the other. Valant [2020] and Gollwitzer et al. [2020] show that areas with a larger share of Republican-leaning voters exhibit less physical distancing, weaker support for stay-at-home orders and more school reopenings. Interestingly, the estimates show that higher vaccination rates are also a strong predictor of higher EIPL.

Second, and importantly, the addition of the different regional attributes increases the predictive power of the regression substantially while reducing the relationship of EIPL with local affluence, education, and parental structure to almost zero. In other words, the negative predictive effect of local affluence, education, and parental structure on EIPL in the above regressions arises primarily because these variables proxy for voter preferences and to a lesser extent vaccination rates, teacher unionization rates and the local wage comparison index. These results offer a new perspective on the nexus between income at the national vs. at the local level and the provision of EIPL.

Third, the inverse relation between EIPL and the share of non-white students is also reduced. Yet, even after controlling for all the state and county-specific attributes, a school with a student body at the 75th percentile of the non-white distribution is predicted to average 3-7% lower EIPL during the 2020-21 school year than a school at the 25th percentile of the distribution.

Fourth, the negative coefficient estimate on test scores and school size remain largely unaffected, indicating that even after controlling all the other school characteristics and regional attributes, school quality and size played an important role for EIPL. By contrast, the regional variables completely absorb the negative association of EIPL with pre-pandemic school spending per student, indicating that school spending did not play a decisive role for EIPL overall but instead picked up systematic differences across regions. Finally, the inverse relation between EIPL and ESSER funding per student becomes somewhat more negative. Hence, even within counties, schools in districts with more ESSER funding per student provided on average not more but slightly less EIPL.

4.5 EIPL for whom? The view from schooling mode trackers

To conclude this section, we assess whether we would reach similar conclusions if we had used the different schooling mode trackers instead of EIPL. We run the same regression as above on school trackers data, focusing on the district or county level to maximize comparability across the results.[18] We summarize here the results and present the full details in the online appendix.

First, the statistical associations between the in-person learning mode of the different trackers and local affluence, education, family structure and race are similar in terms of (most) coefficient signs but are substantially weaker than for our EIPL measure. The weaker association may be explained by the higher level of geography in these data, which removes part of the local variations in those variables. One important exception is that the Burbio coefficient suggests a positive relation between the share of non-white students and in-person learning, while the other trackers indicate a negative correlation (consistent with our findings and those in prior studies).

Second, the schooling mode trackers indicate that ESSER funding per student is associated with addi-

in the 2016 presidential election.

[18]We cannot run the same analysis for EdWeek and exclude the IES-SSD data which is only at the state level. For CSDH, we use the district instead of the school-level data. To control for school characteristics (i.e., grades and school type), we aggregate data to the district or county level, i.e. we measure the share of the student population enrolled in elementary schools, etc.

tional remote learning, except for Burbio and CRPE where ESSER funding per student is associated with fewer weeks of remote learning. The granularity of the EIPL database allows for more precise estimates that help decide between the conflicting results.

Third, the schooling mode trackers also point to a positive association between in-person learning and Republican vote shares during the 2020 election. But once again, there are sizable differences in terms of magnitude across trackers. The schooling mode trackers also show, consistent with our school-level regression, that the mask mandates were negatively related (hence, a substitute policy) to EIPL while the vaccination campaign acted as a complement to school reopenings.

5 Conclusion

We review several school trackers of the options – in-person, hybrid, remote learning – that were more frequently in place during the COVID-19 pandemic, and document a number of systematic differences between them regarding the extent of school closures and reopenings. We then propose a measure of effective in-person learning (EIPL) which we argue allows for a better assessment of the exposure of students in-person learning and its relation to a host of population, school, and regional characteristics. We make the EIPL dataset publicly available for future research on the COVID-19 school disruptions. To conclude, we highlight three questions raised by these data:

1. Why did schools in more affluent and more educated areas with higher funding per student provide less EIPL? We show that this inverse relationship is in large part about political preferences. But why would more Democratic-leaning areas have been more reluctant to let students return to in-person learning? One potential explanation is that independent of political preferences, more affluent and educated parents were on average more likely to be able to work from home and therefore considered the cost of supervising students' virtual learning from home (either in person or by hiring help) more manageable. It might also be that parents had a different perception of the risk of sending students back to in-person school, for instance due to different news and social-media exposure. Both of these explanations contrast, however, with the observation that even within counties, private schools (which generally attract students from wealthier backgrounds) provided more EIPL than public schools. No matter the explanation, it remains that students in more affluent and more educated areas of the U.S. received on average less EIPL.

2. Why did schools with a higher share of non-white students provide less EIPL, even within a given county and controlling for neighborhood poverty and other school characteristics? This striking result defies a simple explanation and yet seems key given the large and persistent educational achievement gaps between students of different races that existed already before the pandemic.

3. Why did schools in districts with more ESSER funding per student not provide more EIPL? One possible reaction is that without ESSER funding, schools would have been closed for even longer. Yet, the absence of a positive relationship arises even within counties and despite controlling for many other school characteristics, which makes this an unlikely explanation. Another potential explanation is that Congress imposed few constraints on how ESSER funding could be used, and according to estimates by Malkus [2021a], less than 20% had been spent by August 2021. If these funds were spent

primarily to improve students' remote learning capacities (e.g. providing students with computers and wireless connections) instead of upgrades to the school buildings and personal protection equipment, then ESSER funding would have primarily facilitated remote learning instead of a return to in-person learning; i.e. its main advertised purpose.[19]

Exploring these questions goes beyond the scope of the paper but they are clearly important to understand the causes and consequences of school closings during the pandemic.

References

Francesco Agostinelli, Matthias Doepke, Giuseppe Sorrenti, and Fabrizio Zilibotti. When the great equalizer shuts down: Schools, peers, and parents in pandemic times. *Journal of Public Economics*, 206:104574, 2022.

Katherine A Auger, Samir S Shah, Troy Richardson, David Hartley, Matthew Hall, Amanda Warniment, Kristen Timmons, Dianna Bosse, Sarah A Ferris, Patrick W Brady, Amanda Schondelmeyer, and Joanna Thomson. Association between statewide school closure and COVID-19 incidence and mortality in the US. *JAMA*, 324(9):859–870, 2020.

Dena Bravata, Jonathan H Cantor, Neeraj Sood, and Christopher M Whaley. Back to school: The effect of school visits during COVID-19 on COVID-19 transmission. *NBER Working Paper 28645*, 2021.

Andrew M Camp and Gema Zamarro. Determinants of ethnic differences in school modality choices during the COVID-19 crisis. *Educational Researcher*, 51(1):6–16, 2022.

Victor Chernozhukov, Hiroyuki Kasahara, and Paul Schrimpf. The association of opening K-12 schools and colleges with the spread of COVID-19 in the United States: County-level panel data analysis. *Proceedings of the National Academy of Sciences*, 118(42), 2021a.

Victor Chernozhukov, Hiroyuki Kasahara, and Paul Schrimpf. Causal impact of masks, policies, behavior on early covid-19 pandemic in the U.S. *Journal of Econometrics*, 220(1):23–62, 2021b.

Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Where is the Land of Opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics*, 129 (4):1553–1662, 2014.

Raj Chetty, John N Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya S. Porter. The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility. *NBER Working Paper 25147*, 2020.

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

[19] The remaining unspent ESSER funds could still be put to good use, for instance by providing summer learning programs for students, especially since ESSER funds were allocated to school districts according to pre-pandemic Title I spending and thus benefited disproportionally schools with students from poorer backgrounds. As pointed out by Malkus [2021a], however, ESSER funding comes with very few constraints and it is unclear to what extent school districts will use the funds for their intended purpose.

Flavio Cunha, James J. Heckman, and Susanne M. Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931, 2010.

Thomas Dee, Elizabeth Huffaker, Cheryl Phillips, and Eric Sagara. The revealed preferences for school reopening: Evidence from public-school disenrollment. *NBER Working Paper 29156*, August 2021.

Emma Dorn, Bryan Hancock, Jimmy Sarakatsannis, and Ellen Viruleg. Covid-19 and education: The lingering effects of unfinished learning. Technical report, McKinsey & Company, 2021a.

Emma Dorn, Bryan Hancock, Jimmy Sarakatsannis, and Ellen Viruleg. Covid-19 and education: The lingering effects of unfinished learning. *Technical report, McKinsey & Company*, 27, 2021b.

EdunomicsLab. National education resource database on schools, November 2021. URL www.edunomics.org/nerds.

Per Engzell, Arun Frey, and Mark D Verhagen. Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, 118(17), 2021.

Zeynep Ertem, Elissa M Schechter-Perkins, Emily Oster, Polly van den Berg, Isabella Epshtein, Nathorn Chaiyakunapruk, Fernando A Wilson, Eli Perencevich, Warren BP Pettey, Westyn Branch-Elliman, et al. The impact of school opening model on SARS-CoV-2 community incidence and mortality. *Nature Medicine*, pages 1–7, 2021.

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.

Nicola Fuchs-Schündeln, Dirk Krueger, Andre Kurmann, Etienne Lale, Alexander Ludwig, and Irina Popova. The fiscal and welfare effects of policy responses to the COVID-19 school closures. *IMF Economic Review*, (forthcoming), June 2022.

Kairon Shayne D Garcia and Benjamin W Cowan. The impact of school and childcare closures on labor market outcomes during the COVID-19 pandemic. *NBER Working Paper 29641*, 2022.

Dan Goldhaber, Scott A Imberman, Katharine O Strunk, Bryant G Hopkins, Nate Brown, Erica Harbatkin, and Tara Kilbride. To what extent does in-person schooling contribute to the spread of Covid-19? Evidence from Michigan and Washington. *Journal of Policy Analysis and Management*, 41(1):318–349, 2022a.

Dan Goldhaber, Thomas J Kane, Andrew McEachin, Emily Morton, Tyler Patterson, and Douglas O Staiger. The consequences of remote and hybrid instruction during the pandemic. *NBER Working Paper 30010*, 2022b.

Anton Gollwitzer, Cameron Martel, William J Brady, Philip Pärnamets, Isaac G Freedman, Eric D Knowles, and Jay J Van Bavel. Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic. *Nature Human Behaviour*, 4(11):1186–1197, 2020.

Clare Halloran, Rebecca Jack, James Okun, and Emily Oster. Pandemic Schooling Mode and Student Test Scores: Evidence from US States. *NBER Working Paper 29497*, November 2021.

Benjamin Hansen, Joseph J Sabia, and Jessamyn Schaller. Schools, job flexibility, and married women's labor supply: Evidence from the COVID-19 pandemic. *NBER Working Paper 29660*, 2022.

Melanie Hanson. Average Cost of Private School. Technical report, EducationData.org, 2021. URL https://educationdata.org/average-cost-of-private-school.

Michael T Hartney and Leslie K Finger. Politics, markets, and pandemics: Public education's response to COVID-19. *Perspectives on Politics*, pages 1–17, 2020.

Youngsoo Jang and Minchul Yum. Aggregate and intergenerational implications of school closures: A quantitative assessment. *Covid Economics, Vetted and Real-Time Papers*, 57:46–93, 2020.

Vladmir Kogan and Stéphane Lavertu. How the COVID-19 pandemic affected student learning in Ohio: Analysis of Spring 2021 Ohio state tests. Technical report, 2021.

Liana Christin Landivar, Leah Ruppanner, Lloyd Rouse, William J Scarborough, and Caitlyn Collins. Research note: School reopenings during the COVID-19 pandemic and implications for gender and racial equity. *Demography*, 59(1):1–12, 2022.

Karyn Lewis, Megan Kuhfeld, Erik Ruzek, and Andrew McEachin. Learning during COVID-19: Reading and math achievement in the 2020-21 school year. Technical report, NWEA Center for School and Student Progress, 2021.

Nat Malkus. How much of federal COVID-19 relief funding for schools will go to COVID-19 relief? Technical report, American Enterprise Institute, 2021a.

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

Tareena Musaddiq, Kevin M Stange, Andrew Bacher-Hicks, and Joshua Goodman. The pandemic's effect on demand for public schools, homeschooling, and private schools. *NBER Working Paper 29262*, 2021.

Zachary Parolin and Emma K Lee. Large socio-economic, geographic and demographic disparities exist in exposure to school closures. *Nature Human Behaviour*, 5(4):522–528, 2021.

Maria J Prados, Gema Zamarro, and Andrew Camp. School reopenings, childcare and labor outcomes during COVID-19. *Unpublished manuscript*, 2021.

Jon Valant. School reopening plans linked to politics rather than public health. *Brown Center Chalkboard, Brookings Institution*, 2020. URL https://www.brookings.edu/blog/brown-center-chalkboard/2020/07/29/school-reopening-plans-linked-to-politics-rather-than-public-health/.

Online Appendix for:
School Closures and Effective In-Person Learning during COVID-19*

André Kurmann
Drexel University

Etienne Lalé
Université du Québec
à Montréal

Contents

*Contact: Kurmann: Drexel University, LeBow College of Business, School of Economics, 3220 Market Street, Philadelphia, PA 19104 (email: kurmann.andre@drexel.edu); Lalé: Université du Québec à Montréal, Department of Economics, C.P. 8888, Succ. centre ville, Montréal (QC), H3C 3P8 Canada (email: lale.etienne@uqam.ca). Aseni Ariyaratne provided excellent research assistance. We thank Dennis Roche from Burbio, Nat Malkus from Return to Learn, and Safegraph for generously sharing their data. We also thank seminar participants at the 2nd joint IZA & Jacobs Center Workshop on "Consequences of COVID-19 for child and youth development" for their comments. All errors are our own.

A Schooling mode trackers

This appendix provides additional details about the schooling mode trackers surveyed in Section 2, and a fuller analysis of the correlation between these data and the population, school, and regional characteristics studied in Section 4 of the paper.

A.1 Data sources

Our review covers eight schooling mode 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.[1]

- 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.[2] 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).

[1] 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.

[2] 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.

A.2 Definition of In-Person / Hybrid / Remote learning

Table A1 complements Table 1 in the main text by reporting the definition of In-Person, Hybrid, and Remote learning implemented in the eight pandemic schooling mode trackers covered by our analysis.

Table A1: 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. 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. For example, if parents can choose between hybrid (2 days per week) or in person (5 days per week), the district is coded as offering in-person instruction." The approach of favoring the most in-person available option may help reduce differences in measured learning modes across the different school trackers.

Second, several school trackers include one or several additional categories to describe the type of learning mode offered to the student population. For instance, CRPE includes the following descriptors "Varies by school", "Varies by grade band", "To be announced", "No information". Education week contains several other descriptors: "No order", "In-person instruction decisions are currently being made on a local level, with states only providing guidelines or recommendations", "States with an order to provide in-person instruction may grant waivers to individual districts", "Closure orders may include exceptions for small groups or particular populations of students". In the CSDH data, there is an additional option for "The school or district was closed during the relevant time period (typically for holidays, spring break, etc.)". Meanwhile, in practice, with exception of EdWeek, the additional categories are almost never used.

Third, and related, the use of additional categories to complement in-person / hybrid / remote learning illustrates well the challenges of describing the complex reality of school modalities during the pandemic. The ESOS definitions (see Table A1) are very telling in this respect. 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. Notice that this is potentially a source of additional noise through frequent changes in the operational status of schools.

A.3 In-person and remote learning: When, where and for whom?

Figures A1a and A1b 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 A1a and A1b 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).[3] Each tracker is used to run several regressions, mirroring our approach in the main text. To appreciate the effects of local affluence, education and family structure, the left-hand side variable is regressed separately against household income, education, or share of double-headed households, together with the share of non-white students and the other variables controls (first three variables reported in each plot). Then, for the share of non-white students and the other variables, all the right-hand side variables are included jointly in the regression. To summarize, each tracker is used to run four regressions in each plot, so that Figures A1a and A1b are both showing the results from 24 different regressions.

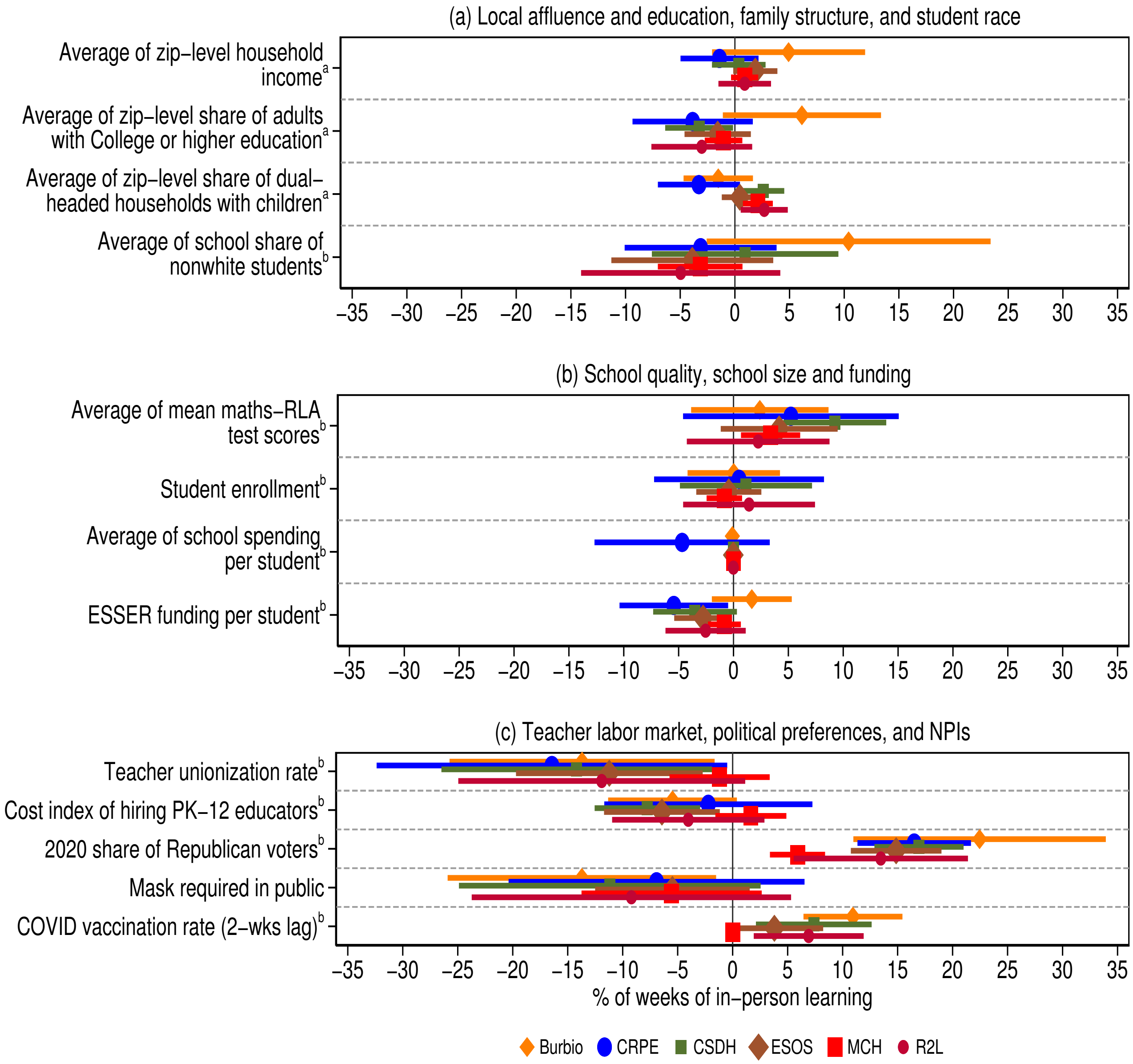
A number of key patterns stand out from Figures A1a and A1b. First, the different school trackers suggest little to no association between in-person or remote learning on the one hand and local affluence, education and family size on the other. This is consistent with the results based on the EIPL dataset 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.[4] 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 and remote learning in Figure A1b, 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 A1a and A1b shows results that, on the whole,

[3] 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 A1a and A1b. For CSDH data, we run the regressions using the school district data as it is more comparable to the other datasets. Thus, the level of geographic analysis in Figures A1a and A1b is either the school district or the county (for Burbio data). Standard errors are clustered accordingly.

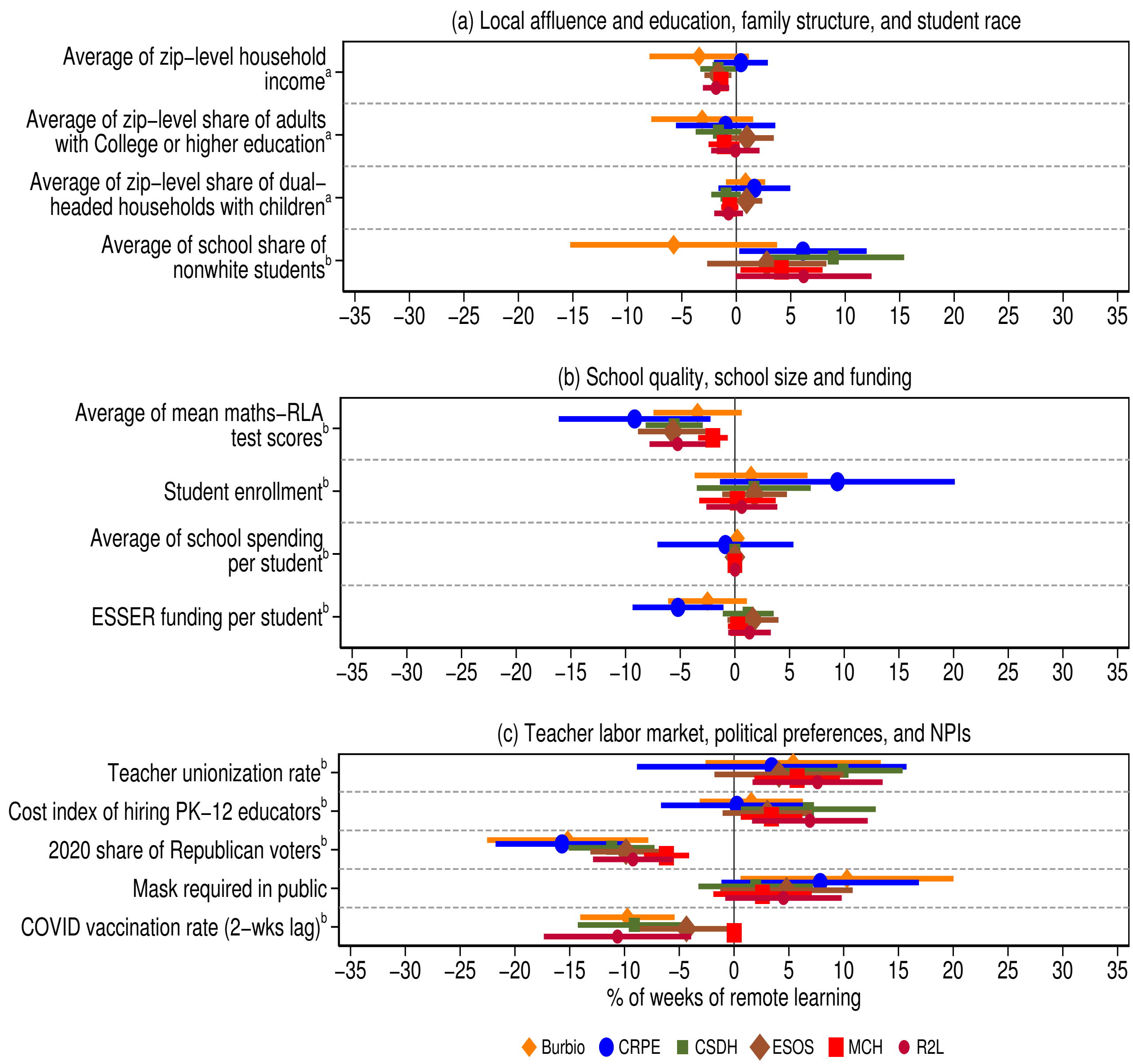
[4] Figures A1a and A1b are also consistent with results based on the EIPL dataset in that, if anything, the relationship between in-person learning and local characteristics is stronger in what concerns local education compared compared to local affluence and family size.

Figure A1a: 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, denoted by ^{a}, 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 denoted by ^{b} 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 so that they show the implied change in remote learning of going from the 25th percentile to the 75th percentile of the distribution of a variable.

Figure A1b: The Relationship of Effective 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, denoted by ^{a}, 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 denoted by ^{b} 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 so that they show the implied change in remote learning of going from the 25th percentile to the 75th percentile of the distribution of a variable.

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. This illustrates well the advantages of the EIPL dataset that allows for a more robust assessment of these correlations.

B EIPL data appendix

This section presents additional information about the NCES data (Subsection B.1); the Safegraph school visits data, our algorithm to match it to NCES datasets, and to construct school weights for the matched data (Subsection B.2); and details on the construction of the EIPL measure (Subsection B.3).

B.1 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 dataset.

The first one is the 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.

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 (see Subsection B.2), and the area-type indicators to assess the relation between EIPL and local school characteristics (see Figure D5).

B.2 Safegraph data, matching algorithm, school weights

B.2.1 Safegraph data 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.[5] 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. Except for a handful of POIs (about 1% of the universe of Safegraph's POIs), each `placekey` comes with

[5] See Safegraph's Places Manual as well as this blog post for additional details.

a 6-digit industry NAICS code.[6] About 80-85% of Safegraph's POIs come with information on visits. 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 visits data, we aggregate up these visits and attribute them to the school that is paired to these non-611110 NAICS POIs.

B.2.2 Matching of Safegraph POIs with NCES school records Our algorithm to match Safegraph's elementary and secondary 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:[7]

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.

   (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.

[6] 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.

[7] 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.

Table B1 describes the outcomes of our merge/match algorithm. We obtain direct merges for about 75,000 schools, which represent more than 60% of schools in the NCES files. Then, through matching based on names and GPS coordinates, we obtain additional matches for about 4,000 schools, and then fuzzy name matches of high quality for about 1,000 schools.[8] We do not include in our analysis fuzzy name matches that do not rely on using schools' names. Likewise, we exclude schools for which the quality of the matching procedure is not high enough.[9] Through this algorithm, we obtain highly reliable matches for 110,644 schools (93,312 public schools and 17,332 private schools).

Table B1: 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 (in %) of schools from the NCES's CCD and PSS files that we either 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 square defined by the 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.

B.2.3 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. Figure B1, which plots the total number of residing Safegraph devices (in counties that contain POIs with NAICS code 611110) over the 2018-2021 period, illustrates the magnitude of these changes.[10] As can be seen, 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.

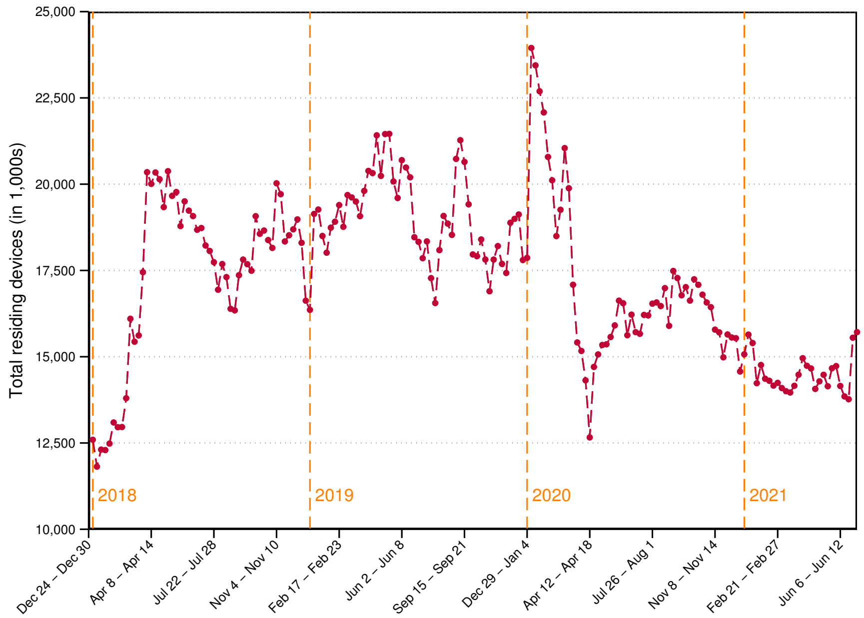
Figure B2 illustrates the impact of these variations on counts of visits to all Safegraph's POIs with

[8]To take one example of fuzzy name matches of high quality, consider Safegraph's "Big Spring Lake Kinderg Sch" described in Footnote 7. 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.

[9]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.

[10]Residing devices are cell phones with a primary nighttime location. Counts of residing devices are released at the CBG level. These counts are typically used for the purpose of normalizing visits data at a granular level.

Figure B1: Safegraph: Number of residing cell phone devices



Notes: The figures show the sum of Safegraph cell phone devices across all counties that contain POIs with NAICS code 611110 ("Elementary and secondary schools").

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%.

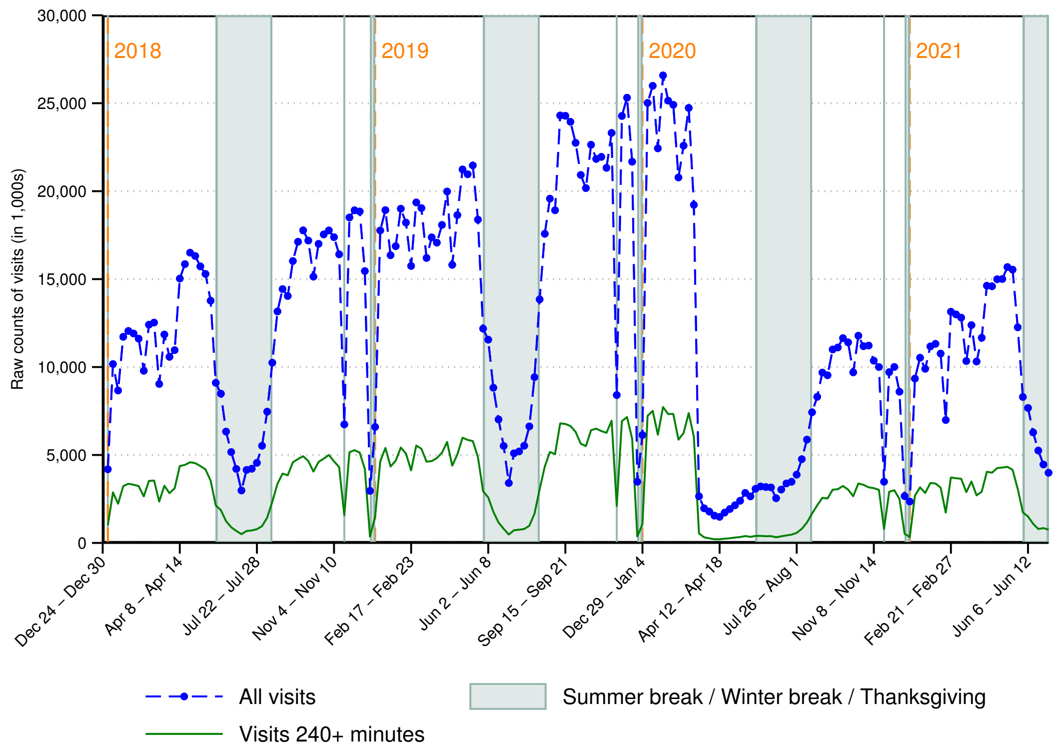
- Then, we drop schools where $\Delta\widetilde{v}_{j,t}$ is larger than 75 more than once, either during the period from beginning of September 2019 to November 2019 or the period from beginning of September 2020 to the end of the sample period. This procedure intends to purge the data from extreme 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.

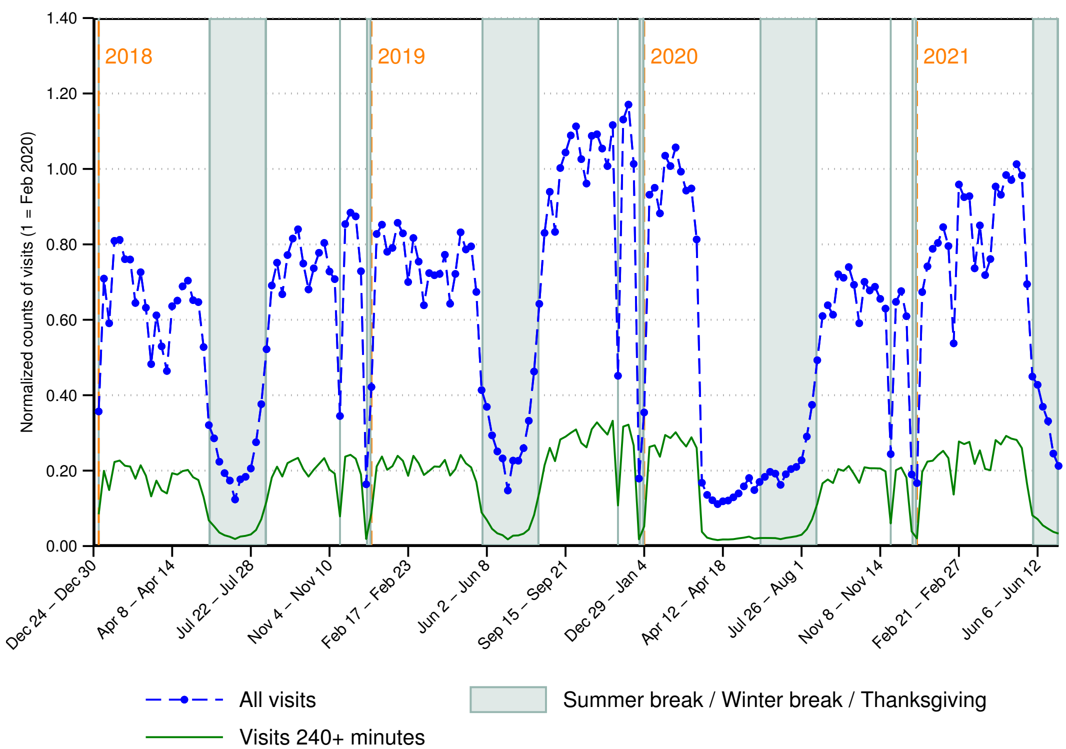
B.2.4 School weights As explained in Section 3 of the main text and Subsections B.2.2 and B.2.3 of this appendix, the dataset of our analysis includes only a subset of the schools from the pooled CCD-PSS

Figure B2: 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").

file after filtering out schools with sparse or noisy visit data. We augment the dataset with school-level weights to alleviate concerns about its representativeness. 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 $\hat{\Pr}\{y_j = 1\}$.[11] 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 Table B2).

B.2.5 Comparison of selected sample to the NCES universe of schools Table B2 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.

Through direct merges, matching and fuzzy-name matching (see Appendix B), we combine the CCD-PSS file to Safegraph data. Out of 123,234 schools in the CCD-PSS file, we obtain 110,644 matches to the Safegraph data for which the quality of the match is high enough to be considered as reliable. We call this subset of schools the "matched" sample. Out of those schools, we retain 73,194 schools in our analysis and discard the remaining ones that suffer from excessively sparse or noisy $\Delta\widetilde{v}_{j,t}$ (see Subsection B.2.3 for details). The remaining schools constitute the "in scope" dataset. Table B3 compares different observables in the full CCD-PSS file (columns 1 and 4) with the schools that we match to Safegraph data (columns 2 and 5) and to the subset of these schools that we retain in our analysis of school visits (columns 3 and 6). The characteristics of the full CCD-PSS file and our matched dataset are very similar, suggesting that the algorithm does not selectively pick certain schools while discarding others. In columns 3 and 6 all the entries (except sample counts) are computed using the school weights described in Subsection B.2.4, which by construction make the in-scope dataset representative of the universe of CCD-PSS schools.

B.2.6 A closer look at the school visits data As a final step towards preparing our combined Safegraph-NCES "in scope" dataset, 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 $\Delta\widetilde{v}_{j,t+1}$. This adjustment implements the assumption that during the school year 2020-21, schools did not reopen for only one week at a time.

Figure B3 plots the distribution of average changes in visits $\Delta\widetilde{v}_{j,t}$ 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. Despite the various adjustments (Section 2 of main text and Subsection B.2), we see a substantial variation in school visits: each panel in Figure B3 uses 4 weeks of data for each school j, and yet a non-trivial share of changes in visits fall outside of the [-20%, +20%] interval. This said, some of this dispersion may capture variations in school activity across months. For instance, some schools may not reopen right in the beginning of September 2019, which would explain why the

[11] 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 $\hat{\Pr}\{y_j = 1\}$.

Table B2: Comparison to the NCES digest of education's statistics

Number of educational institutions		
	NCES table 105.50	CCD & PSS
	(1)	**(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	CCD & PSS
	(1)	**(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	CCD & PSS
	(1)	**(2)**
Public Schools	3,170	2,911
Private Schools	482	401
All	**3,652**	**3,312**

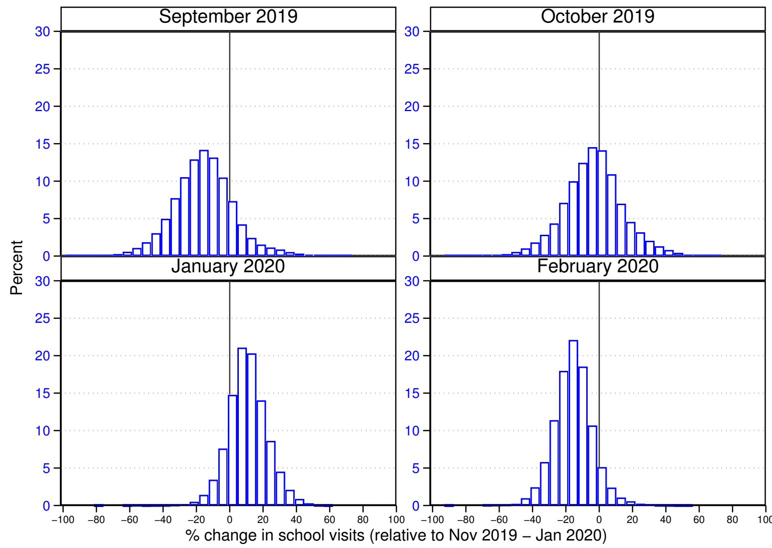
Notes: NCES numbers refer to the year 2017-2018 (most recent release of the NCES's tables covering both public and private schools for the same school year). Public schools classified as "Other", denoted by [a], includes special education, alternative, and other public schools not classified by grade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

Table B3: Comparison between all schools and schools from the in-scope dataset

	Public schools			Private schools		
	All	Matched^a	In scope^b	All	Matched^a	In scope^b
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(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 data and with visits data that is neither too sparse or too noisy (see Subsection B.2.3). Except for the sample count, all the statistics for the "In scope" data are based on the school weights computed in Subsection B.2.4. % 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.

Figure B3: Distribution of changes in school visits before and during the base period



Notes: The figures show the distribution of the average change in school visits during 4 months prior to the pandemic.

distribution is shifted to the left. Similarly, in January 2020, some schools might start later than others and therefore, the distribution is also shifted to the left.

Figure B4 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.[12]

The middle and lower panels of Figure B4 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 B4 shows a clearer recovery in school visits, though with substantial mass around 0 or higher.

B.3 Details on construction of EIPL measure

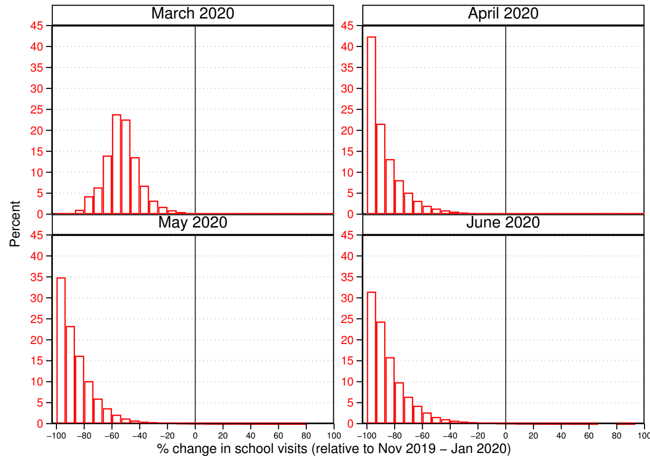
Our main methodological contribution is to construct a measure of EIPL that exploits the information contained in school visits from Safegraph and learning modes from Burbio and Return2Learn. As explained in Section 3, the analysis revolves around the estimation of Equation (4), which is repeated here for reference:

$$T_{c,t} = \alpha + \beta\Delta\widetilde{v}_{c,t} - \gamma H_{c,t} + \varepsilon_{c,t}. \tag{B.1}$$

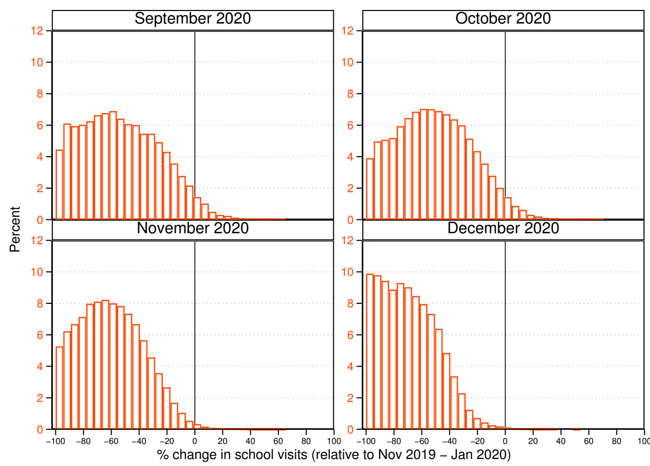
[12] In April, May and June in the upper panel of Figure B4, we observe some schools with changes in visits not lower than -60% or -80%. To understand how this relates to the upper map of county-level loss of EIPL (Figure 1, main text), recall that: 1) these changes in visits are translated into EIPL 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 B4: Distribution of changes in school visits during the pandemic

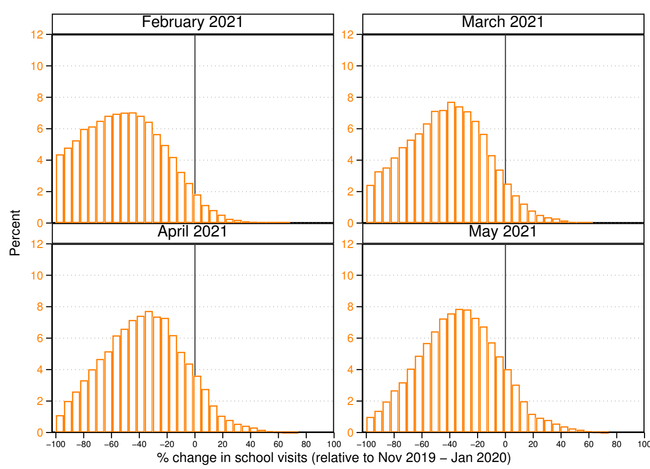
(a) March 2020 to May 2020



(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.

$T_{c,t}$ and $H_{c,t}$ are shares of traditional and hybrid learning modes in county c during week t, $\Delta\widetilde{v}_{c,t}$ are changes in school visits aggregated to the county-level using student enrollment in public schools, α, β, γ are the regression coefficients, and $\varepsilon_{c,t}$ is the regression residual. When working with R2L data, instead of the county level we work with data at the school-district level. Essentially, Equation (B.1) configures a panel regression that can be estimated at different levels of geographic aggregation and/or over different sample periods.[13] We focus on the Core-Based Statistical Areas (CBSA) and state levels and proceed through the following steps:[14]

1. Restrict the data to counties or school districts for which our Safegraph-NCES dataset includes at least 5 schools.

2. For each geographical level (CBSA, state) and each school tracker (Burbio, R2L), find the sample period to estimate Equation (B.1) that yields the highest R-squared. In practice, we find that the best fitting sample periods are mostly during the Fall semester, in the sense that less than 25% of all regressions have a sample period that extends beyond mid-February 2021.

3. For each CBSA and each school tracker (Burbio, R2L), use the best fitting sample period to estimate Equation (B.1) and obtain two sets of estimates, $\left\{\hat{\beta}_{\text{Burbio-CBSA}}, \hat{\gamma}_{\text{Burbio-CBSA}}\right\}$ and $\left\{\hat{\beta}_{\text{R2L-CBSA}}, \hat{\gamma}_{\text{R2L-CBSA}}\right\}$. Similarly, obtain state-level sets of estimates $\left\{\hat{\beta}_{\text{Burbio-state}}, \hat{\gamma}_{\text{Burbio-state}}\right\}$ and $\left\{\hat{\beta}_{\text{R2L-state}}, \hat{\gamma}_{\text{R2L-state}}\right\}$. Store the R-squared that are associated with each set of estimates.

   (a) Some CBSAs have empty estimates for either Burbio or R2L. This occurs when we have insufficient variation in $T_{c,t}$ or $H_{c,t}$, or when $T_{c,t}$ or $H_{c,t}$ are perfectly negatively related to each other; see Footnote 14.

4. At this point we have (at best) four sets of estimates to choose from to determine $\hat{\beta}$ and $\hat{\gamma}$ for each school j. Rank the estimates based on the associated R-squared, i.e. denote by $\hat{\beta}_{(k)}$ and $\hat{\gamma}_{(k)}$ the ranked estimates, where $\hat{\beta}_{(4)}$ and $\hat{\gamma}_{(4)}$ refer to the highest and $\hat{\beta}_{(1)}$ and $\hat{\gamma}_{(1)}$ to the lowest R-squared estimates. Set $\hat{\beta} = \hat{\beta}_{(k)}$ where k is the largest index such that $0 < \hat{\gamma}_{(k)} < 1$.

The fact that our procedure can be applied in almost all instances is due to a combination of factors. First, our Safegraph-NCES dataset has extensive coverage of schools for which $\Delta\widetilde{v}_{j,t}$ is neither too noisy or sparse. Second, by exploiting information from both Burbio or R2L, we increase the number of sets of estimates from which to choose from in the last step of our procedure. 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.

There are, however, a few exceptions. For all four set of estimates (after finding the best fitting sample periods), the R-squared turn out to be no higher than 0.25 for most areas in Arkansas and Maine. This reflects low quality of the Burbio and R2L data in those areas. For instance, in Arkansas the regressions with R2L data at both the CBSA and state levels and the CBSA-level regressions with Burbio data almost always suffer from insufficient variation. As a result, we are left with no option but to use the state-level Burbio estimates, which suffer from a R-squared close to 0. In Maine, the best fitting R-squared is 0.23, except in one of the CBSA. To address these shortcomings, we replace the $\hat{\beta}$ and $\hat{\gamma}$ for schools in those

[13] When the sample period used to estimate Equation (B.1) 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.

[14] Obviously Equation (B.1) can also be run directly at the county or school district levels. The downside of this approach is that some counties or school districts either have no time-variation in $T_{c,t}$ or $H_{c,t}$, or $T_{c,t}$ and $H_{c,t}$ almost perfectly negatively related to each other. The latter occurs when $V_{c,t} = 0$ for most t, making traditional learning mechanically related to hybrid learning; i.e. $T_{c,t} \approx 100 - H_{c,t}$. In a regression context, this implies $\gamma \to 1$ and $\beta \to 0$ since $\Delta\widetilde{v}_{c,t}$ is subject to idiosyncratic variation. As a result, steps 2 and 3 of our approach do not work well at these finer geographical levels. We sometimes run into a similar issue at the CBSA level.

areas of Arkansas and Maine with Burbio estimates from neighboring states (states of the West South Central division for Arkansas, states of the New England division for Maine). Meanwhile, in both states, we do have a few CBSA-level estimates that we do not override given that they have a R-squared higher than 0.25.[15] Although not a validation of our approach, these CBSA-level $\hat{\beta}$'s turn out to be not too different from the estimates that we construct for Arkansas and Maine using neighboring states.

C Details on data used in regressions

ACS data The socio-demographic and income variables (mean household income, share of individuals with some College or higher educational attainment, share of dual-headed households with children) 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). We aggregate data using the so-called 'total ZIP ratio', which is the ratio of all addresses for each Census tract (the first eleven digits of the CBG code) associated with each zip code. To measure population density, we combine the ACS population estimates with 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/). From the USDA, we also use the Rural-Urban continuum codes (see https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx).

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. "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 for methodological details. We use additional data from EDGE in robustness checks presented in Appendix D. We use the school neighborhood poverty estimates, constructed by EDGE based on data from the Census Bureau's American Community Survey which allow 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.

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 information from 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). The NCES numbers, which measure the percentage share of teachers in a union or employees' association, are based on the public school teacher data files of the Schools and Staffing Survey (SASS), whose microdata is not publicly available.

[15] In Maine, we have CBSA-level estimates for the Portland-South Portland-Biddeford metropolitan area. In Arkansas, we have CBSA-level estimates for Fayetteville-Springdale-Rogers and Memphis metropolitan areas, but note that both are multi-state metropolitan areas.

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). SEDA is a data initiative launched in 2016 with the purpose of providing 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 (see Appendix D for additional information). The data we use are from version 4.1 of the SEDA data, which is the most recent version as of this writing.

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/). It builds on the federal Every Student Succeeds Act (ESSA) passed in December 2015 which, among other provisions, stipulates that states must report for every public school (and local educational agency) the total per-pupil spending of federal, state and local money disaggregated by source of funds for the preceding fiscal year. In practice, the school spending data tends to be scattered across different states' website, but NERD$ gathers these data together. The data we use are from the latest 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, that is to say the funds from the Coronavirus Aid, Relief, and Economic Security (CARES) act, the Coronavirus Response and Relief Supplemental Appropriations (CRRSA) act and the American Rescue Plan (ARP). Data are available at the level of school districts, and the R2L database comes 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 COVID 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). Data on COVID vaccinations at the county level are daily rates from the CDC. We download these data from the COVID from the Opportunity Insights Economic Tracker repository (see https://github.com/OpportunityInsights/EconomicTracker). To aggregate each variable 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 of county-level health data 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. State-level data on voter turnout rates are taken from the U.S. elections project (see http://www.electproject.org/). We use turnout rates for the 2018 and 2020 general elections.

NPIs data Data at the county-week level on Non-Pharmaceutical Interventions (NPIs) are downloaded from the repository of the Centers for Disease Control and Prevention (CDC; see https://data.

cdc.gov/). We use information on 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 county-week time series are from the September 10, 2021 update of the CDC data.

OA data In robustness checks presented in Appendix D, 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. The variables from OA that we use are: (i) 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; (ii) 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; (iii) the average rent for two-bedroom apartments in 2015. For (i) and (ii), we use data pooled across genders and races/ethnicities.

D Additional tables and figures

This section presents complementary analyzes of the distribution of EIPL across regions and time (Subsection D.1); of the relation between EIPL and school types, grades and schools' locality (Subsection D.2); additional regression results behind the results presented in Section 5 (Subsection D.3); and descriptive statistics for the main regression variables (Subsection D.4).

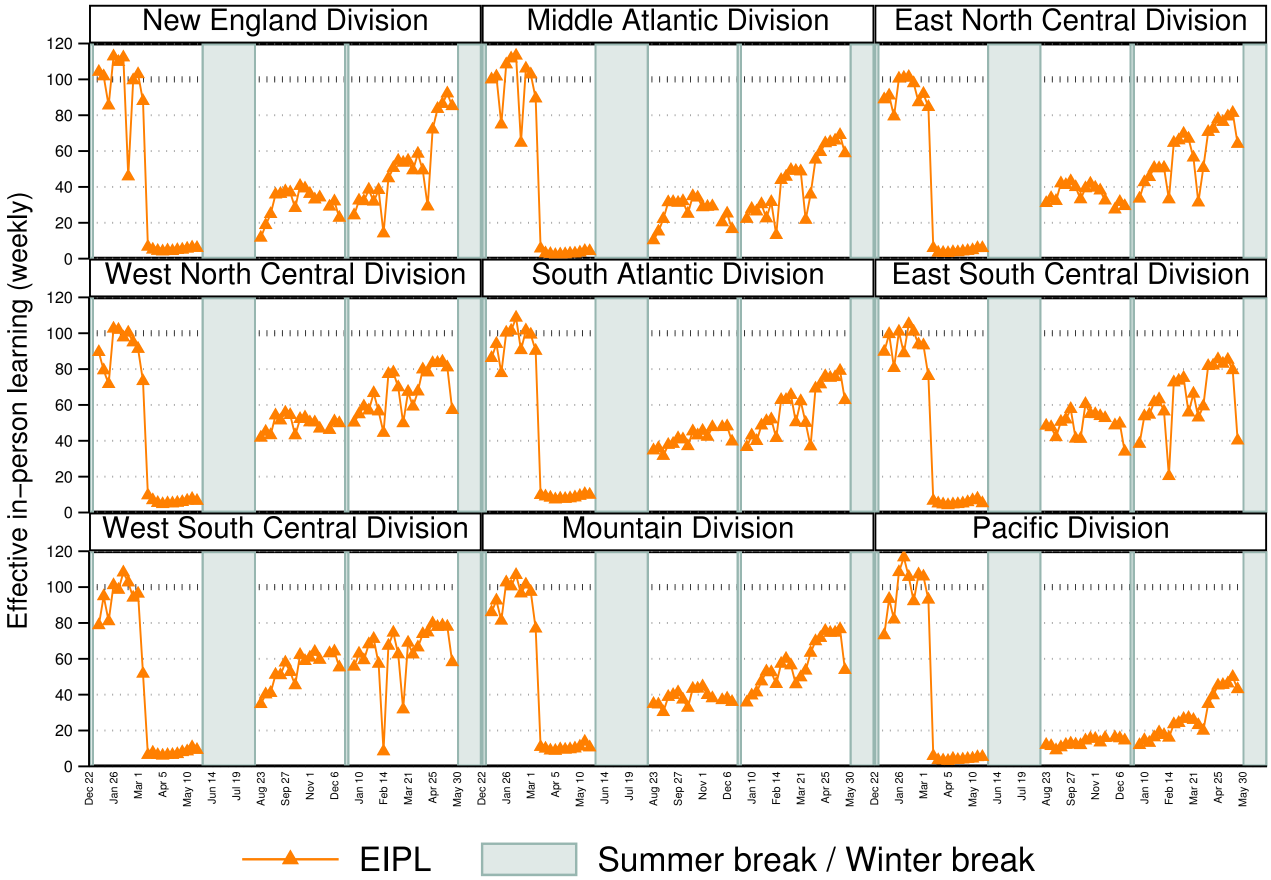
D.1 Regional disparities in EIPL over time

Figure D1 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.

The upper panel of Figure D2 reinforces the finding from Figure D1, 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 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

Figure D1: 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. U.S. Census Divisions are: 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).

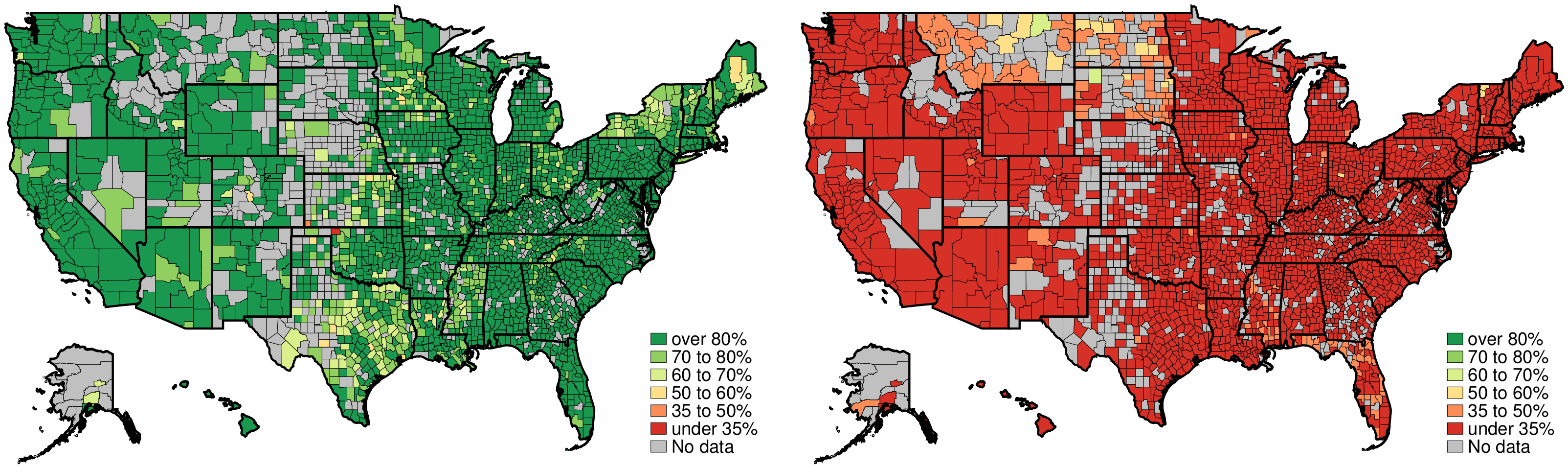
reported in Figure D3 by showing back-to-school dates for each of the nine Census divisions, reveals major differences across different parts of the country. For example, in the East South and West South Central divisions, students usually head back to school at the beginning of August, whereas in New England and in the Middle Atlantic division back-to-school dates are typically after Labor Day.[16,17] We use this information (available at the county-level) in the lower panel of Figure D2 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

[16] 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.

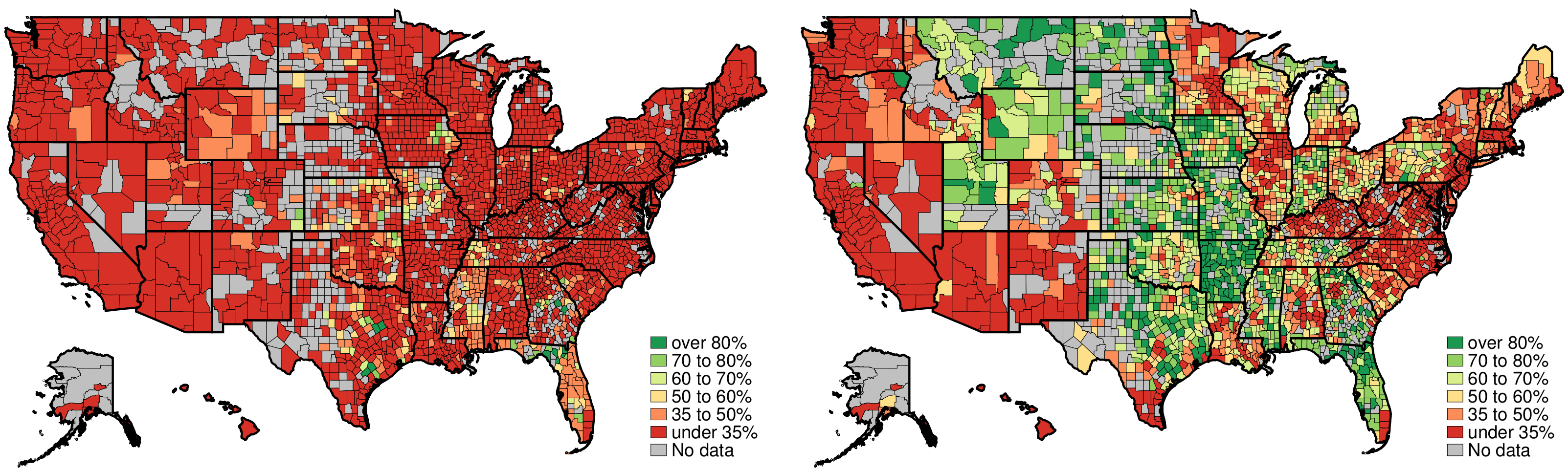
[17] In Figure D1, the shaded area denoting the Summer break of 2020 covers the weeks from May 31st until August 22nd. For many schools across the country, this time interval is only an approximation of the Summer break since, as shown in Figure D3, back-to-school dates are not uniform across regions.

Figure D2: Effective in-person learning across U.S. counties

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

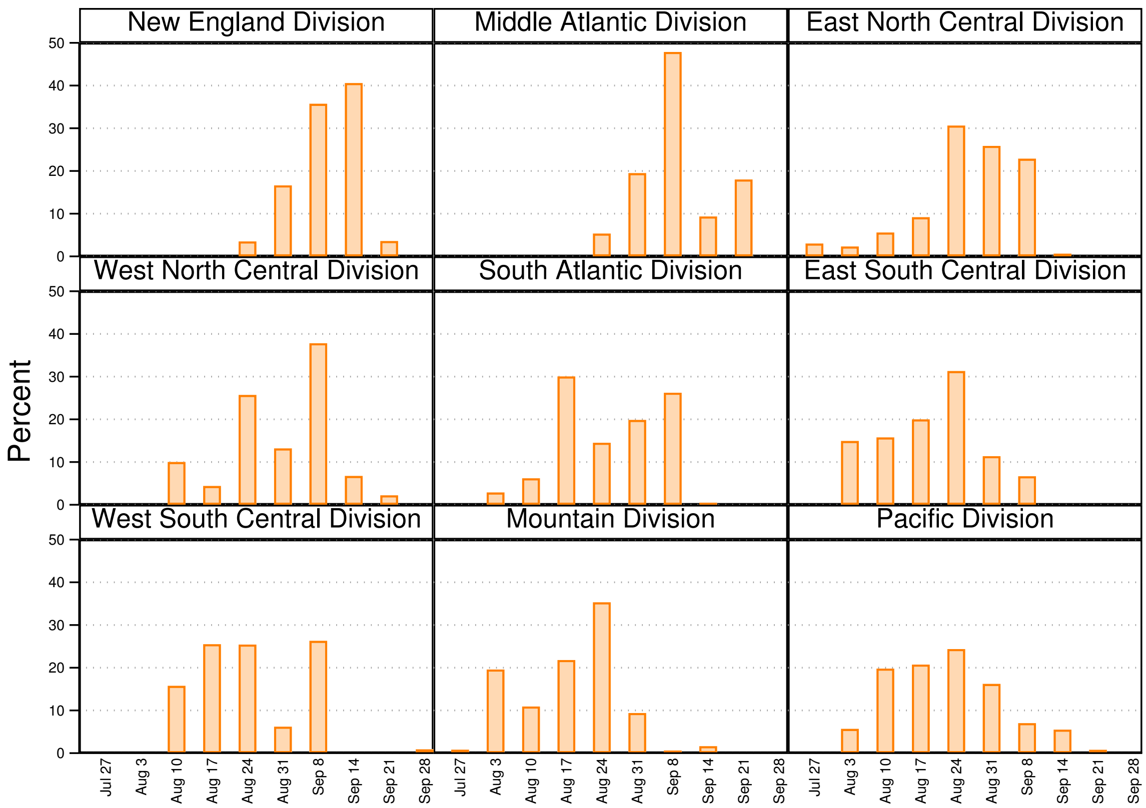


(b) 4 weeks before (left) and after (right) back to school dates according to pre-COVID school district calendars



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 D3: Back to school dates (based on pre-COVID school district calendars), by Census divisions



Notes: The figure shows the percent of public schools in each Census division that would open on the days in question according to pre-COVID school district calendars. U.S. Census Divisions are: 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).

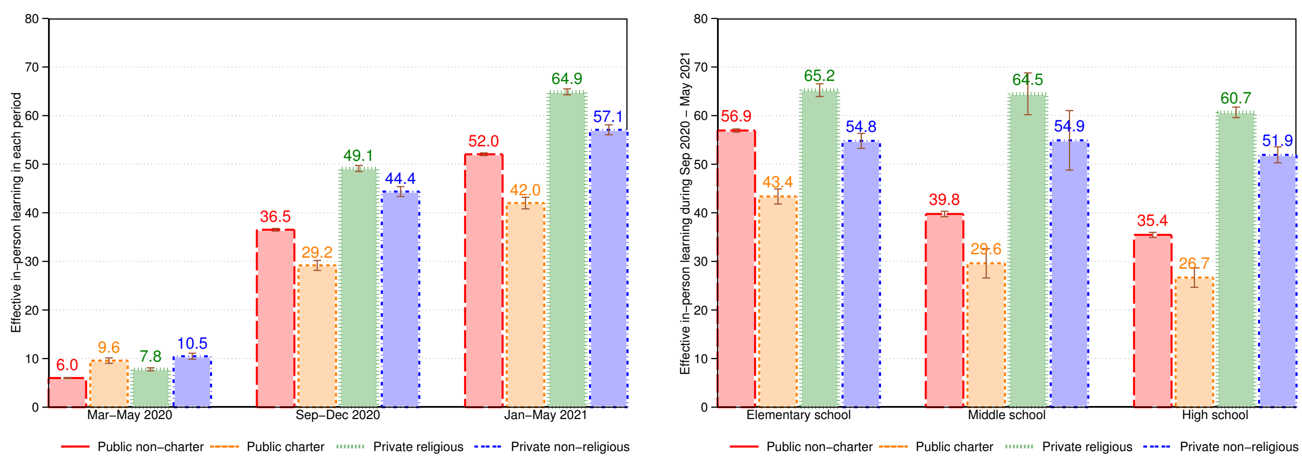
well measured by Burbio, (ii) reopening of private schools that may be asynchronous with that of public schools, (iii) schools reopening for in-person learning earlier than what they usually do in normal times. Then, on the right-hand side of the panel, we observe many counties shifting from close-to-zero 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 D2 before and after usual back-to-school dates.

D.2 Relation of EIPL with school type, grade and locality

Panel (a) of Figure D4 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.

Figure D4: 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.

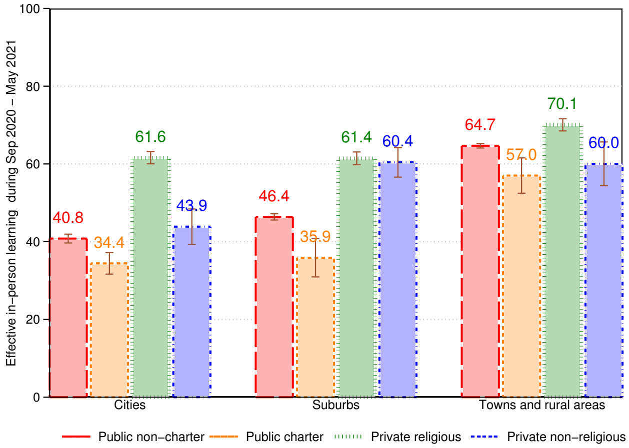
Panel (b) of Figure D4 reports on differences in average EIPL between September 2020 and May 2021 by school type and school grade. Across all four school types, EIPL is highest for elementary schools and lowest for high schools. For private schools, the difference in EIPL across school grades is smaller than for public schools.[18] 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.

Another potentially important school characteristics to understand EIPL during the pandemic is the type of surroundings in which schools are located. Schools that are located within a city have on average lower EIPL than schools in suburbs, which themselves have lower average EIPL than schools located in a town or a rural area. This relationship is robust in the sense that, as shown in Figure D5, it holds true for each school type (public non-charter, public charter, private religious, and private non-religious) 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 D5 also shows that the magnitude of the EIPL gap between public and private schools

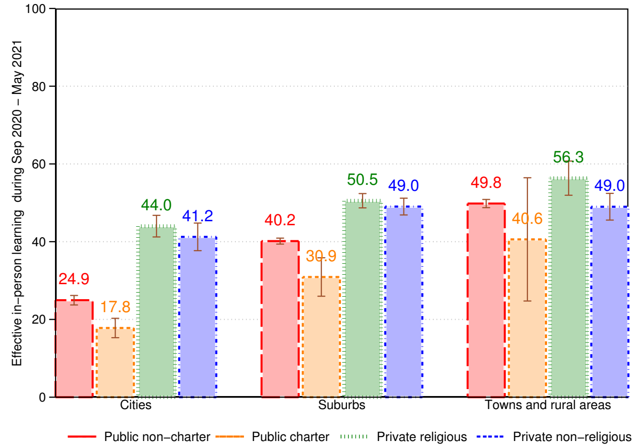
[18] Note that for elementary schools, EIPL is slightly higher for public non-charter schools than for private non-religious schools. This change in ranking of school types compared to the ranking across all school grades is due to geographical differences in the relative prevalence of private non-religious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

Figure D5: Effective in-person learning by school type, locale and U.S. region

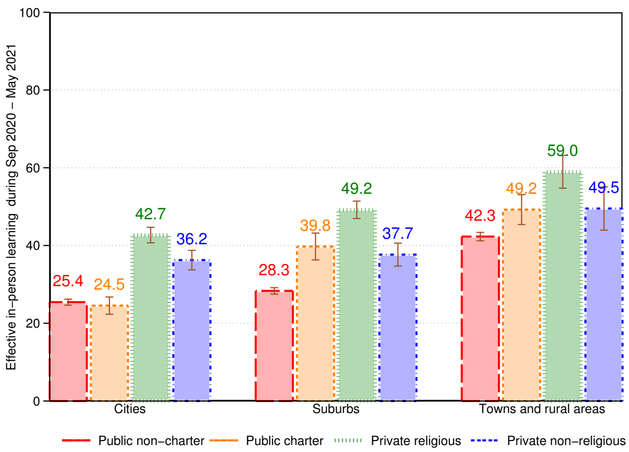
(a) Midwest



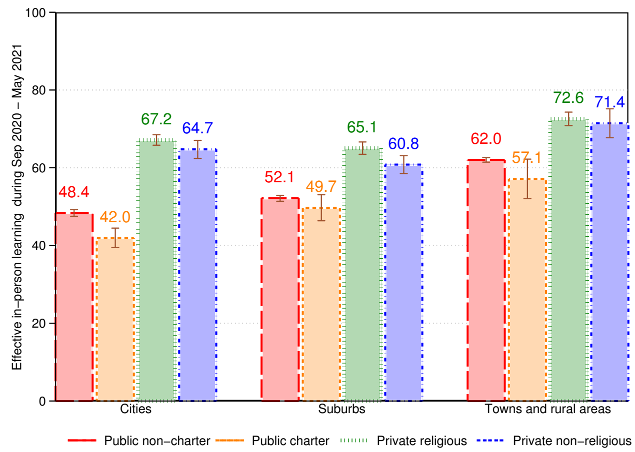
(b) Northeast



(c) West



(d) South



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

differs across regions; for instance it is larger in cities of the Northeast region of the country. Notice that the relations of EIPL with school type, grade and locality in Figure D5 represent unconditional averages, and therefore do not control for other predictors that light up in the OLS regressions.

D.3 Additional regression results for Section 4

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

We begin in Table D1 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. Adding these controls does not change the results noticeably, and the estimates on these controls are in line with the results shown in Figure D4. 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; see Subsection D.4 for additional information.

Table D1: 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 Appendix B.2. 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.

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 D2 repeats the results from Table D1 above as a reference.[19]; 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.

Table D3 analyzes the role of geography. Column (1) in panel (a) repeats the final regression in Table D2 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.

[19] The estimates for household income, share of college educated and school neighborhood poverty are exactly as in columns (1) - (3) of Table D1. The estimate for share of non-white students is slightly different because this estimate is obtained while controlling jointly for all three measures.

Table D2: 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 Appendix B.2. 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 D3: 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 Appendix B.2. 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.

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 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.[20]

Panel (b) of Table D3 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.

Regression results presented in Figure 4 Table D4 presents the estimates reported in Figure 4 of the main text. In order to understand better the role of the county-level regressors, Table D5, 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, as in Column (3) of Table D4. Foremost, Table D5 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)).

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;[21] restricting the sample to counties with at least 10 public schools, which reduces the

[20] 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.

[21] 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

Table D4: Regression results for Figure 4

Dependent variable	Effective in-person learning (EIPL)		
	(1)	**(2)**	**(3)**
Zip-level household income[a]	-5.37***	-5.27***	0.02
	(0.48)	(0.58)	(0.33)
Zip-level share of College educated[a]	-7.87***	-8.51***	-2.51***
	(0.45)	(0.43)	(0.41)
Zip-level share of dual-headed households[a]	-3.16***	-2.92***	-0.56
	(0.70)	(0.74)	(0.37)
School share of non-white students[b]	-20.29***	-15.27***	-6.18***
	(1.60)	(1.34)	(0.86)
Mean maths-RLA test scores[b]		5.12***	4.55***
		(0.81)	(0.60)
Student enrollment[b]		-3.56***	-2.92***
		(0.28)	(0.23)
School spending per student[b]		-4.96***	-0.87**
		(0.68)	(0.37)
ESSER funding per student[b]		-0.52	-0.98*
		(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	0.13	0.15	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 B.2. 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 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 D5: 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 Appendix B.2. 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 (see Appendix C).

sample size almost threefold, leave the results mainly unchanged.

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 D6, 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 Table D4 for reference, while columns (2), (4) and (6) show the estimates for the different races/ethnicities; Panel (b) performs a similar analysis for private schools. Table D6 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.

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 (see Subsection C). In Table D7, we show effects of

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.

Table D6: 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 Appendix B.2. 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 C).

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 Columns (2) and (3) of D4 for reference. 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.

Table D7: 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 Appendix B.2.4. 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 (see Appendix C).

Other indicators of local affluence Table D8 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. Columns (1), (3), (5) and (7) run a quasi-univariate regression, where only the

Table D8: 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 Appendix B.2. 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. See Appendix C for details.

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 D1. Columns (2), (4), (6) and (8) of Table D8 add all the covariates included in our main regression, and are therefore comparable to results shown in Table D4.

Table D8 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 D8 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.

D.4 Description of school-level regression variables

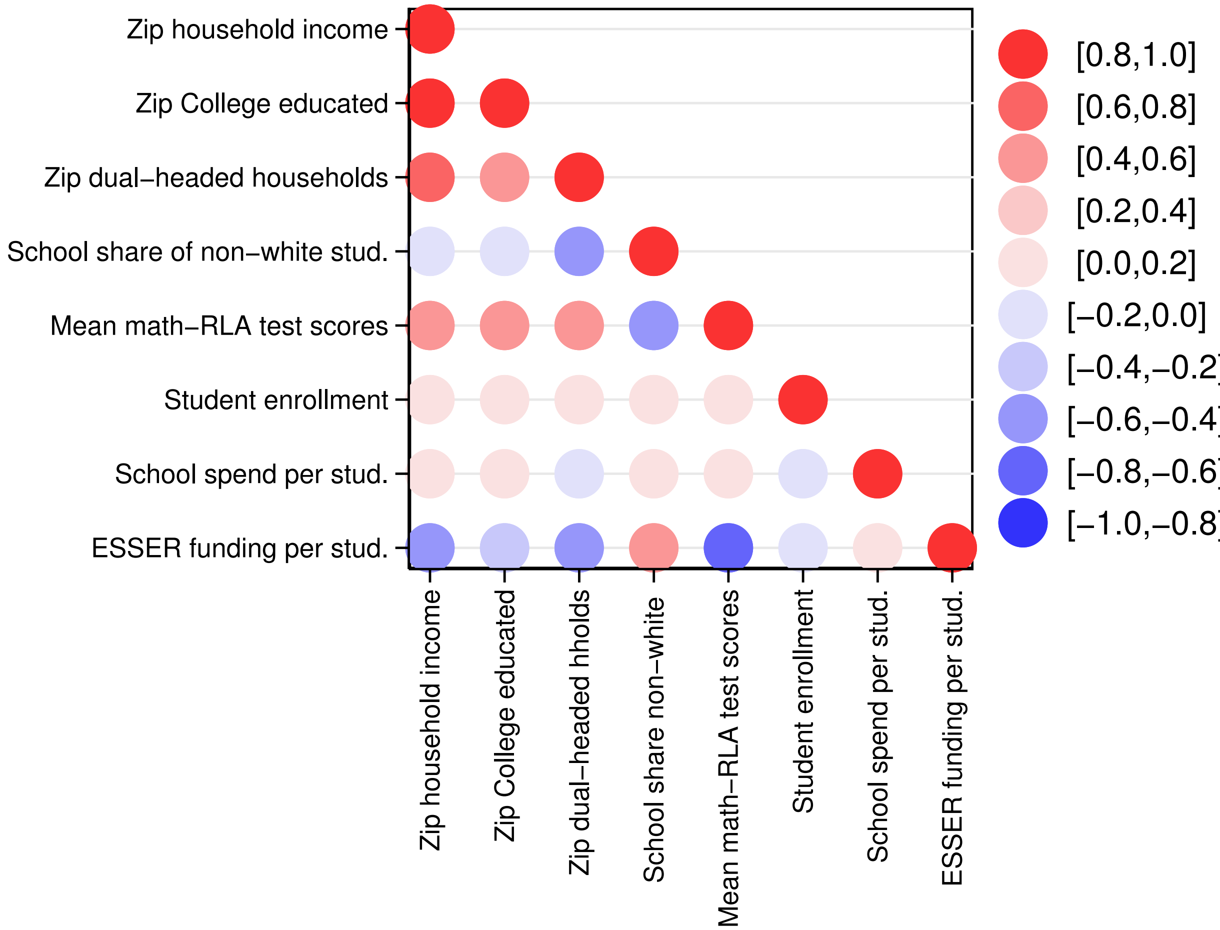
In Section 4, we measure the effects of local and school/district variables on EIPL by looking at the effects of going from the 25th to the 75th percentile of the distribution of a variable. Table D9 presents descriptive statistics for the sample of public and private schools of our analysis. In particular, observe that private schools are on average located in more affluent areas, have a lower proportion of non-white students and are much smaller than public schools in terms of student enrollment.

Table D9: 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 with the school weights calculated as explained in Appendix B.2.

Figure D6: 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 with the school weights calculated as explained in Appendix B.2.

To complement panel (a) of Table D9, Figure D6 presents the correlations between the regressors. As the top left corner shows, the first two affluence measures (household income and share of adults with College or higher education) are highly correlated with each other, with a correlation coefficient of 0.83. The share of dual-headed households with children is also positively related with these measures: the correlation coefficients are respectively at 0.60 and 0.47. The school share of non-white students is negatively related to the affluence measures. The correlations range from -0.47 (correlation with the share of dual-headed households with children) to -0.05 (correlation with zip-level share of adults with College or higher education). The other interesting correlations in this figure are those between ESSER funding per student and the affluence measures. The correlations are negative and within the -0.59 to -0.35 range. Since the share of non-white students is negatively related to the affluence measures, it is positively correlated with ESSER funding per student (correlation of 0.44). There is also a strong inverse relation between ESSER funding per student and pre-COVID district level test scores, with a coefficient correlation of -0.69.

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.