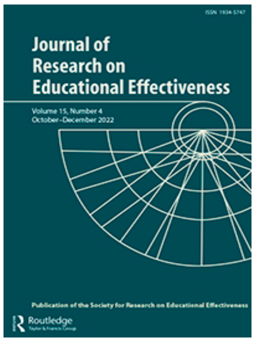

Routledge
Taylor & Francis Group

Journal of Research on Educational Effectiveness

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uree20

Consistency between Household and County Measures of Onsite Schooling during the COVID-19 Pandemic

Carly Lupton-Smith, Elena Badillo Goicoechea, Megan Collins, Justin Lessler, M. Kate Grabowski & Elizabeth A. Stuart

JOURNAL OF RESEARCH ON EDUCATIONAL EFFECTIVENESS
https://doi.org/10.1080/19345747.2022.2131660

METHODOLOGICAL STUDIES

Consistency between Household and County Measures of Onsite Schooling during the COVID-19 Pandemic

Carly Lupton-Smith[a], Elena Badillo Goicoechea[b], Megan Collins[c], Justin Lessler[d], M. Kate Grabowski[e], and Elizabeth A. Stuart[f]

[a]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; [b]Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; [c]Wilmer Eye Institute, Johns Hopkins School of Medicine and Berman Institute of Bioethics, Johns Hopkins University, Baltimore, Maryland, USA; [d]Department of Epidemiology, University of North Carolina Gillings School of Public Health, Chapel Hill, North Carolina, USA; [e]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; [f]Departments of Biostatistics and Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

ABSTRACT

The academic, socioemotional, and health impacts of school policies throughout the COVID-19 pandemic have been a source of many questions that require accurate information about the extent of onsite schooling occurring. This article investigates school operational status datasets during the pandemic, comparing (1) self-report data collected nationally on the household level through a Facebook-based survey, (2) county-level school policy data, and (3) a school-level closure status dataset based on phone GPS tracking. The percentage of any onsite instruction within states and counties are compared across datasets from December 2020 to May 2021. Sources were relatively consistent at the state level and for large counties, but key differences were revealed between units of measurement, showing differences between policy and household decisions surrounding children's schooling experiences. The consistency levels across sources support the usage of each of the school policy sources to answer questions about the educational experiences, factors, and impacts related to K-12 education across the nation during the pandemic, but it remains vital to think critically as to which unit of measurement is most relevant to targeted research questions.

ARTICLE HISTORY
Received 24 March 2021
Revised 5 July 2022
Accepted 17 August 2022

KEYWORDS
Virtual learning; COVID-19 pandemic; educational outcomes; local policy; measure concordance

Introduction

Going into the fall of 2020 in the midst of the COVID-19 pandemic, schools across the United States had to make a decision about how to educate children and adolescents effectively and safely. Schools varied in their instructional policies from fully online, to hybrid online and onsite, to all onsite, and these policies often changed multiple times during the course of the pandemic (Honein et al., 2021). Furthermore, decision-making

CONTACT Carly Lupton-Smith clupton1@jhu.edu Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 615 N Wolfe Street, Baltimore, MD 21205, USA.

occurred at various levels: county, district, school, or even household, as schools often gave parents the choice about whether to send their children in for onsite instruction. The potential implications of these instructional models on children, their families, and their communities are substantial. However, measuring the extent of onsite instruction is complex; how it is done, and the quality of the underlying data, can have notable effects on the conclusions drawn from studies that examine links between school operations and relevant outcomes during and beyond the COVID-19 pandemic.

There are reasons to be very interested and concerned regarding links between school instructional models during the pandemic and COVID-19, social, mental health, and academic outcomes. On the one hand, when schools were closed in March of 2020, evidence indicated that these closures were correlated with lower levels of COVID-19 spread (Auger et al., 2020) and there have been multiple papers since aiming to study the links between onsite schooling and COVID-19 transmission (e.g., Falk et al., 2021; Goldhaber et al., 2020; Harris et al., 2021; Honein et al., 2021; Lessler et al., 2021; Zimmerman et al., 2021), with varying definitions of school "reopening" (i.e. all onsite, hybrid) and consideration of mitigation measures such as mask wearing and physical distancing. For example, Lessler et al. found that individuals living with a child attending school onsite between November 2020 and February 2021 had an increased risk of COVID-19, but this risk was reduced if mitigation measures were employed in the school (Lessler et al., 2021). Similarly, during the summer of 2021, schools with no mask requirement in two counties in Arizona were three and a half times more likely to have a school-associated COVID-19 outbreak compared to schools that had a mask requirement starting at the beginning of the academic year (Jehn et al., 2021).

Beyond COVID-19 infection implications of school instructional modalities, there are also potential large impacts on children and their families in other ways. The shift to virtual learning during the early months of the pandemic and onward will likely affect student outcomes going forward, and these effects could be worse for individuals who have faced inequities in the education system already (Hamilton & Ercikan, 2022). Food insecurity can increase for families who count on school as a supplier of meals for their children (McLoughlin et al., 2020; Van Lancker & Parolin, 2020), and schools are often the sites of routine health care for children, such as vision and hearing screenings and even vaccinations. Previous research has shown, for this group of students, that summer vacations lead to a loss of academic achievement compared to students with higher socioeconomic status (Van Lancker & Parolin, 2020). There are reasons to expect a similar dynamic during pandemic-related school closures (Eyles et al., 2020). In the Netherlands, learning loss occurred even with a short lockdown, and this loss was up to 60% larger among students who lived in less educated homes (Engzell et al., 2021). Virtual instruction is made more difficult in households that have problems with computer and internet access, which is more common in less educated households as well (Bansak & Starr, 2021). Further research found that students with low academic achievement and learning motivation before the pandemic began had a more difficult time dealing with virtual learning (Berger et al., 2021), and that student adaptability impacts academic emotion (Zhang et al., 2020) and self-efficacy (Martin et al., 2021) in the online learning environment. Thus, these school closures have potential to exacerbate disparities across areas and households in the country (Armitage & Nellums, 2020).

There are therefore many factors that school leaders have to consider when making school policy decisions during this pandemic, including the instructional modality and also any preventive measures put in place for onsite instruction. And until Centers for Disease Control and Prevention guidance in February 2021, there was little formal guidance from the federal government on school reopening, leading to substantial variability in these decisions across the country (Center for Disease Control, 2021). Limited work has been done so far to understand these processes and decisions. One study found that mass partisanship and the strength of the teacher's union were the strongest predictors of a district's decision to reopen or not reopen schools (Hartney & Finger, 2022). Interestingly, the severity of COVID-19 in the area was a potential predictor in their model but was not identified as having a strong relationship.

A challenge in studying the links between school operational status and key outcomes of interest is the variability in operations across the country and the lack of standardized data on the topic. A few sources exist (very few on truly national data), and they vary in how they collect data, how many areas in the country they report on, how often they update these reports, how they define operational status, and to which level of measurement (i.e., district, county) they aggregate. In addition, most sources are measured only at the district or county level and do not indicate individual household decisions about onsite instruction. Since many districts offer families a choice to stay fully virtual during the COVID-19 pandemic even if there are onsite options, it is crucial to obtain individual household information about schooling behaviors in order to fully understand the extent of onsite instruction and how it varies across districts (Harris et al., 2021). It is also important to understand how well that household self-report relates to district-, county-, and state-level policies.

In this article, we thus compare regional-level instructional status information with a large-scale national daily survey containing household-level responses about schooling behaviors. Measuring schooling experiences during the pandemic has been especially challenging because there is no "gold-standard," or external benchmark that represents the true amount of onsite versus online schooling throughout the country. Instead, we must obtain information about schooling experiences through comparing the data that does exist. This situation is increasingly common in education, in which a set of imperfect measures available through administrative or other (e.g., social media) datasets are all that are accessible to represent key factors worthy of study. Therefore, we must develop techniques to learn about phenomena of interest through the data that is available and must determine how to best consider consistency—or lack thereof—across sources. This article is a case study of such a setting. The comparisons in this article show the similarities and differences between datasets that assess school policies at different levels (i.e., household, district, county, and state). Although we cannot effectively conclude which dataset is the most accurate, each source described in the following sections provides a different lens from which we can view schooling policies and experiences. Ultimately, if the estimates from each level of analysis are relatively consistent, researchers can elect to use whichever source or sources best address the research questions of interest in future studies. Estimates could differ depending on the level of analysis (i.e., household, school, district, county), but the hope is that all are capturing the same underlying factors—examining their consistency thus can help reveal how valid any of them might be.

Research questions for this study included: (1) According to each source, what was the nationwide percentage of any onsite school from December 2020 to May 2021, and what were the percentages for each county?, (2) How consistent were monthly state-level and county-level aggregated percentages of any onsite school across sources?, and (3) How do the differences in county-level monthly onsite school percentage reported by each data source relate to county-level factors like income inequality, population density, and unemployment rate?

Data and Methods

This section details the schooling data available from four national datasets, along with measures used from these datasets and others. Tables 1 and 2 provide details of the sources used to compare data on a state- and county-level (Table 1) and on a district-level (Table 2), including the specific questions used to assess instructional modalities.

COVID-19 Trends and Impact Survey

The Delphi Group at Carnegie Mellon University U.S. COVID-19 Trends and Impact Survey, in partnership with Facebook (CTIS) is a daily cross-sectional survey that stratifies by U.S. state and invites a new stratified random sample of adult Facebook users to take the survey each day by displaying the survey invite at the top of a user's News Feed (Barkay et al., 2020; Salomon et al., 2021). The current study uses the U.S. data, which began collection on April 6, 2020, and has had approximately 50,000 respondents every day since. The survey includes questions about COVID-19 symptoms, testing, and diagnosis; health-related behaviors such as mask wearing; schooling; and mental health and financial worries (Kreuter, 2020; Salomon et al., 2021). Each respondent's Federal Information Processing Standard (FIPS) county code and zip code are obtained for geographical mapping and regional aggregation. The survey includes weights that reflect distributions at the state level and adjust for coverage and non-response biases (Salomon et al., 2021). Specifically, Facebook calculates weights using a two-stage process that involves first weighting by an inverse propensity score to account for non-response bias, and then post-stratification weighting to adjust the Facebook user age and gender distribution to that of the general population (Barkay et al., 2020). A recent comparison between the CTIS data and the 2019 American Community Survey (ACS) found that the weighted survey sample was comparable to the ACS in terms of age and geography but overrepresented women and individuals with more than a high school education (Salomon et al., 2021). We performed analyses using this dataset under a Data Use Agreement between our research group and the Carnegie Mellon University Delphi Group.

Questions regarding PreK-12 students' educational experiences were added in Wave 5 of the survey, which began on November 24, 2020 (Delphi Group, 2020). Data used in this study were thus from December 1, 2020, to May 31, 2021, with data between December 23, 2020 and January 2, 2021, removed under the assumption that most

Table 1. Datasets used in state- and county-level comparisons.

Source	Update frequency	Unit of measurement	Measure of onsite schooling	Dates	Data collection	Question	Sample size	Weighting
Delphi U.S. CTIS[a]	Daily	Household	If children in household with children in 1st–12th grade are attending any onsite instruction	12/1/20 – 5/31/21 (excluding 12/23/20 – 1/2/21)	Facebook platform-based survey	Do any of the following apply to any children in your household (pre-K–grade 12)? Going to in-person classes full-time; Going to in-person classes part-time	1,277,837 respondents from all states and FIPS codes	Facebook weights
Burbio[b]	Weekly	County	Percentage of county offering onsite instruction	Weeks of 12/4/20 – 5/28/21 (excluding last two weeks in December)	Searches Internet for information on 1,200 districts; aggregates them to the county level, imputing for smaller counties	Percentage of county offering: only virtual, combination, or traditional in-person schooling	All counties	County population
U.S. School Closure and Distance Learning Database[c]	Monthly	School	Percentage of schools open for onsite instruction	Months of December 2020 - May 2021	SafeGraph GPS data on 80,785 public schools in the country	Decline in total visits to school compared to the same month in 2019 At least 50% decline → "closed" Report the percentage of schools in the county that are closed	80,785 public schools	District enrollment (for state-level comparisons)

Note: Overall comparison is on the monthly percentage of any in-person schooling in the state/county; All three sources cover nearly every United States county and cover all U.S. states.
[a]https://delphi.cmu.edu/covidcast/survey-results/.
[b]https://cai.burbio.com/school-opening-tracker/.
[c]https://osf.io/tpwqf/.

Table 2. Comparison of 200 districts, January 2021.

Source	Update frequency	Data collection	Question
Burbio[a]	Weekly	Gathers information through web search on top 200 largest districts	Percentage of district offering: only virtual, combination, or traditional in-person
U.S. School Closure and Distance Learning Database[b]	Monthly	SafeGraph GPS data on 80,785 public schools in the country	Decline in total visits to school compared to the same month in 2019 At least 50% decline → "closed" Report the percentage of schools in the district that are closed
MCH[c]	Variable	Reached out to individual districts to get information	Teaching method: on premises, hybrid, online, unknown, other, or pending

[a]https://cai.burbio.com/school-opening-tracker/.
[b]https://osf.io/tpwqf/.
[c]https://www.mchdata.com/covid19/schoolclosings.

schools were on winter break. Data from Puerto Rico and anywhere outside of the United States were removed. Analyses of the CTIS were restricted to participants who responded that they had children in their household in 1st through 12th grades. Importantly for the comparisons below, the type of school that children attended was not obtained, so children could have attended public, private, or other types of schools. The full sample had 1,277,837 respondents with a non-missing FIPS code. This information is summarized in Table 1.

For comparison with other data sources, we summarized household-level measures from the CTIS at the state- and county-level using the weights provided by the survey (Barkay et al., 2020). For comparison at the county level, the sample included only survey participants from the 460 most populous counties in the country to allow for more accurate comparison with other data sources, as smaller counties had very small sample sizes from the CTIS. Therefore, the data for county-level comparisons consisted of 909,829 total respondents. CTIS was not in any district-level comparisons because participants were not asked to provide information about their school district so data could not be aggregated to that level. The Delphi Group maintains an overview of the CTIS data and current aggregated results (Delphi Group, 2022), and for a current publication providing a description of the data, see Salomon et al. (2021).

Burbio

Burbio is a data service company that reports information about schools, government, library, and community events, including schooling related factors during the COVID-19 pandemic. Burbio reports publicly available weekly measures for the percentages of public schools in each county in the country offering only virtual school, a combination of virtual and onsite school, or only onsite school (see Table 1) (Burbio, 2020a). Burbio's data collection began in September 2020, with more systematic data collection starting in December of 2020, so we use their weekly data from December 2020 to May 2021, removing the last week of December due to holiday breaks from school. To obtain data, Burbio searches school district websites or uses other sources like Facebook to gather information on 1,200 mostly large districts that represent 46% of U.S. K-12 public schools (Oster et al., 2021). If a district offers multiple options for schooling (i.e., onsite or hybrid), that district is classified using the more in-person option (i.e., onsite) (Burbio, 2020a; Oster et al., 2021).

Burbio ultimately releases aggregate data at the county level, weighting by rough enrollment of the districts based on whether the districts are large, medium, or small (Burbio, personal communication, January 11, 2021; Harris et al., 2021). They estimate data for the smallest counties by imputing from nearby counties (Harris et al., 2021) and update policies for each county weekly. For state-level comparisons, we calculated a weighted average of Burbio county-level schooling policy percentages, weighting by county population. For county-level comparisons, we used the most populous 460 counties because these counties had the most direct estimates (Burbio, 2020a). Finally, Burbio also releases detailed weekly information about the 200 largest districts, so we obtained the schooling percentages for these districts in January 2021 for district-level comparisons with the next two sources described. Available resources discuss Burbio's methodology (Burbio, 2020a), aggregated results (Burbio, 2020b), and a publication using these data (Oster et al., 2021).

U.S. School Closure and Distance Learning Database

The U.S. School Closure and Distance Learning Database (SCDL) uses anonymous phone GPS data from SafeGraph to estimate year-over-year changes in the number of visits to a school in each month of 2020 and 2021 compared to the same month in 2019 (Parolin & Lee, 2021). This database includes phone data from 109,905 public and private schools in 94% of school districts and 98% of counties, and it reports the change in school visit percentages at the district-, county-, census tract-, and state-level for public schools only ($n = 80{,}785$, 82% of all public schools in United States).

The SCDL currently has monthly data from January 2020 to June 2021, and we used their monthly percentages of schools closed per county and per state from December 2020 to May 2021 in our analyses. We also obtained their district-level percentages for the largest 200 districts in January 2021 for a brief district-level comparison. The codebook and publicly available data for the SCDL are available online (Parolin & Lee, 2022), and a publication explaining and utilizing these data was written by Parolin and Lee (2021).

MCH School District Operational Status Data

One final dataset was only part of district-level comparisons and is less of a focus of this work but provides additional information. This dataset comes from MCH Strategic Data (2020). The MCH school district operational status dataset contains district-wide designations of teaching method, which could be on premises, hybrid, online, unknown, other, or pending (MCH Strategic Data, 2020). To collect data, the MCH team reached out to each district across the country (Harris et al., 2021). District teaching methods are not updated at regular time increments (whereas CTIS is updated daily, Burbio weekly, and SCDL monthly), so we used MCH data from the January 31, 2021, update (see Table 2). We removed districts marked as private (n = 60), Catholic (n = 16), or missing (n = 11) for these comparisons and used only the largest 200 public school districts to align with the Burbio district-level data. Because the MCH data were not updated at regular time increments and because of the difficulty in aggregating from district to county or vice versa, we did not use the MCH data in county- or state-level comparisons. MCH data are available publicly online (MCH Strategic Data, 2020), and a publication using it was written by Harris et al. (2021).

Other Sources

There are a handful of other datasets that document school policies and household schooling behaviors during the pandemic, including the U.S. Census Bureau PULSE survey (U.S. Census Bureau, 2021), EdWeek (DATA: State Dashboards on COVID-19 in Schools and Instructional Models, 2020), UAS survey (University of Southern California, 2020), CRPE (CRPE, 2020), NCES School Pulse Panel (Institute of Education Sciences, 2022), Return to Learn (R2L) Tracker (AEI, 2022), COVID-19 School Data Hub (CSDH) (Oster et al., 2021), and JHU eSchool+ (Johns Hopkins Berman Institute of Bioethics, 2020). We did not include these sources for comparison because some only included a select number of districts across the country (EdWeek and CRPE), questions were not easily comparable with the question of interest for this study (PULSE and JHU), the time frame did not match up with our analysis time frame (NCES), or data could not be as easily aggregated to the county-level (UAS, R2L, CSDH).

Measures

Each dataset described above measures onsite schooling percentages in a unique way, as detailed in the "Measure of Onsite Schooling" column in Table 1. In the CTIS, participants with school-aged children were asked if any of the following applied to children in their household (1st-12th grade): going to onsite classes full-time, going to onsite classes part-time (Delphi Group, 2020). We grouped these responses into a category of "any onsite school," which was the case if an individual said any children were either going to onsite classes full-time or part-time, or "no onsite school" if both responses were marked "No." Results aggregated to the county-level and state-level therefore are the Facebook survey-weighted percentages of households with school-aged children in the county/state who had children attending any onsite school. Estimates from Burbio were already at the county-level, so these estimates are percentages of each county in the country offering any onsite school (including part-time). We aggregated these estimates to the state-level as well (weighted by county population) and used district-level percentages of onsite school for a brief comparison among large districts.

The SCDL includes a measure of "the estimated share of schools with at least a 50% year-over-year decline in in-person visits in a month" in a region, which they denote as the percentage of schools that are closed or mostly closed in that month and region. This 50% cutoff is recommended by the SCDL researchers, after checking with several other datasets, including the Household Pulse Survey (U.S. Census Bureau, 2021) and Education Week's district data (DATA: State Dashboards on COVID-19 in Schools and

Instructional Models, 2020), which yielded high levels of consistency. (Less consistency was found with these other sources when using a cutoff of 25% or 75%; Parolin & Lee, 2021). We subtract this percentage from 100% to estimate the percentage of schools that are open for onsite schooling in a given month and region. Finally, the MCH school district operational status dataset estimate provides district-wide designations of teaching method, which could be on premises, hybrid, online, unknown, other, or pending. This estimate is only used in a brief district-level comparison.

Aside from the measures of onsite schooling, we collected county characteristics to estimate associations with county percentage of onsite schooling and differences in onsite schooling reported across sources. County factors included: monthly unemployment rates (U.S. Bureau of Labor Statistics, 2021), population density from 2010 (Ykzeng, 2020), parental educational attainment from 2014 to 2018 (USDA ERS, 2021), monthly COVID-19 cases and deaths (USA Facts, 2021), and 2018 census data on race breakdown, population, health insurance, Gini coefficient[1] (an indicator of income inequality) (Sitthiyot & Holasut, 2020), essential employment, computer access, and percentages of residents working at home (Glenn, 2019; U.S. Census Bureau 2020a, b). Finally, we pulled state populations from World Population Review (2021) for weighted summary statistics.

Analyses

Preliminary analyses involved exploratory graphs and correlations to assess the consistency between the three datasets at the county and state levels. We assessed consistency primarily for the percentage of students attending any onsite school, which included fully onsite and hybrid (partially onsite). We compared these variables on the state- and county-level by month from December 2020 to May 2021 for the CTIS, Burbio, and SCDL sources, and secondary district-level comparisons in January 2021 were performed using the Burbio, SCDL, and MCH data. We also calculated county-level pairwise differences between the sources for the percentage of students with any onsite school and described these differences through descriptive statistics and histograms.

Additionally, we merged county-level factors with the school policy data to investigate the relationships between county-level factors and percentages of any onsite schooling across sources for the most populous 460 counties in the United States. Based on correlations between these factors and the percentages of any onsite instruction according to each source, some county factors were selected to be included in a multiple linear regression to answer the question: how do the differences in county-level monthly onsite school percentage reported by each data source relate to county-level factors? The equation for the regression is shown below (1). Here, Y_{ijk} represents the percentage of onsite school for county i (ranging from 1 to 460) in month j (December through May) from source k (CTIS, Burbio or SCDL). $Source_k$ indicates which source the percentage came

[1]The calculation of the Gini coefficient is based upon the Lorenz curve, which compares the cumulative normalized rank of income with the cumulative normalized income. From there, the Gini coefficient is the ratio of the area between this Lorenz curve and the line indicating perfect equality (where the cumulative normalized rank of income is equal to the cumulative normalized income), divided by the area under the line indicating perfect equality (Sitthiyot & Holasut, 2020).

from; $Month_j$ indicates which month the percentage came from; and X_{ij} is a matrix of monthly county-level values from census data of the following variables: percentage of White residents, percentage on public assistance, percentage working at home, Gini coefficient, population density, percentage of essential employees, percentage with computers, percentage of adults with a high school degree, unemployment rate, and COVID-19 cases normalized by county population. A random intercept allowing for variation by county is denoted by u_i. Focal coefficients of this analysis are the interaction terms, which indicate if the size of the differences between data sources are related to county-level factors.

$$Y_{ijk} = \beta_0 + u_i + \beta_1 * Source_k + \beta_2 * Month_j + \beta_3 * X_{ij} + \beta_4 * Source_k * Month_j + \beta_5 * Source_k * X_{ij} + \epsilon_{ijk},\ where\ \epsilon_{ijk} \sim N(0, \sigma^2),\ u_i \sim N(0, \tau^2),\ and\ cov(u_i, \epsilon_{ijk}) = 0. \tag{1}$$

Results

Concurrence of State-level Estimates across Data Sources

Across the country between December 2020 to May 2021, an average of 57.5% of households with children in first through twelfth grade had children receiving any onsite instruction according to the Delphi U.S. CTIS. When weighting by district enrollment, 58.5% of students in the nation attended a school offering onsite instruction according to the U.S. SCDL Database. And in terms of the amount of onsite schooling offered, the percentage of students living in counties with schools offering onsite instruction

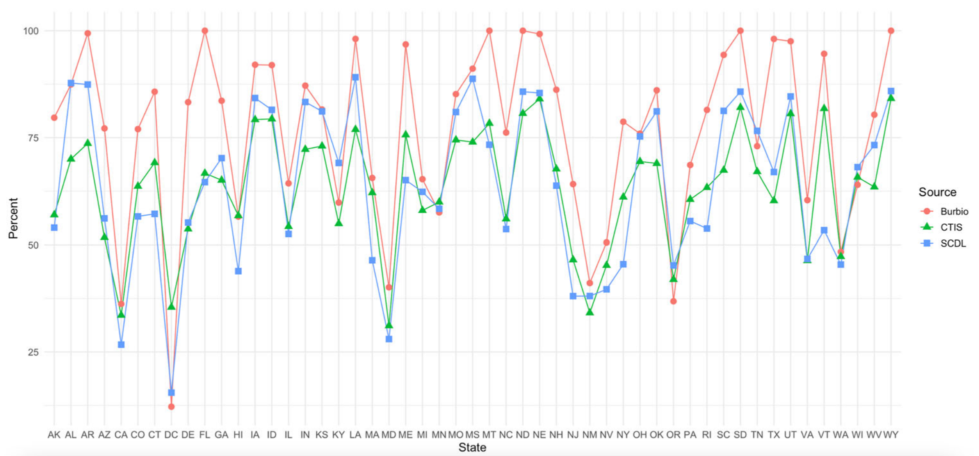

Figure 1. Percentage of students participating in any onsite instruction by state: December 2020–May 2021. Burbio reports weekly county-level percentages of onsite school. To calculate state-level Burbio percentages, we weighted by county population to get statewide percentages of onsite school and then averaged across all time. CTIS reports individual responses of whether their child is attending any onsite school. To calculate state-level CTIS percentages, we weighted individual responses by the survey-provided weights. SCDL reports percentage of schools in the district that were open for any onsite school. To calculate state-level SCDL percentages, we weighted district percentages by district enrollment.

Table 3. Correlations and RMSD between sources of measures of the state-level percentages of any onsite schooling.[a]

Any onsite school Correlation (RMSD)	Burbio	CTIS	SCDL
Burbio	1 (0)	0.89 (21.33)	0.80 (22.04)
CTIS	0.89 (21.33)	1 (0)	0.87 (11.03)
SCDL	0.80 (22.04)	0.87 (11.03)	1 (0)

[a]Total sample size is n = 306 (51 states * 6 months).

Table 4. Descriptive statistics of largest 460 counties and all counties in the country.

	Top counties[a] mean (SD)	All counties[b] mean (SD)
Percentage of any onsite school: CTIS	58.8 (19.8)	71.2 (23)
Percentage of any onsite school: Burbio	75.7 (36.0)	84.5 (31.2)
Percentage of any onsite school: SCDL	60.3 (25)	77.4 (25.2)
Population	521,193 (730,064.3)	105,569.6 (333,984.4)
Percentage of Black	12.4 (12.8)	9.1 (14.5)
Median household income	64,698.1 (16,654.4)	51,540.9 (13,683.3)
Percentage on public assistance	2.3 (1.2)	2.3 (1.5)
Gini coefficient	0.5 (0)	0.4 (0)
Population per square mile	1,290.6 (4,313)	266.3 (1,747.4)
Percentage with no health insurance	8.6 (4)	10 (5)
Percentage of essential employees	27.3 (3.7)	33.1 (7.2)
Percentage of households with computer	90 (3.6)	83.4 (6.9)
Percentage of adults with high school degree	89.2 (5)	86.5 (6.2)
Monthly unemployment rate	5.9 (2.1)	5.4 (2.2)
Monthly COVID cases	5,313.6 (13,100)	1,026.5 (5,335)
Monthly COVID deaths	80.6 (204.4)	16.9 (83)

[a]Total sample size is 18,852.
[b]Total sample size is n = 2,760 (460 counties * 6 months).

according to Burbio was 72.6%. Figure 1 shows the reported percentages by state and data source, demonstrating fairly strong consistency across states during the time period considered. The correlations between the monthly, state percentages of any onsite school for the three sources are high (Table 3), indicating high levels of consistency between the estimated percentage of any onsite school across states. Specifically, the CTIS monthly state percentages of households with children receiving any onsite instruction have a 0.89 ($p < .001$) correlation with Burbio percentages of students attending schools offering onsite instruction and 0.87 ($p < .001$) with SCDL percentages of students attending schools offering onsite instruction, and Burbio and SCDL are correlated at 0.80 ($p < .001$). The root mean squared differences (RMSD) between sources are relatively large, though. The pairwise RMSDs are: 21.33 between CTIS and Burbio, 22.04 between SCDL and Burbio, and 11.03 between CTIS and SCDL. Therefore, the state-level estimates are highly correlated, but Burbio has a large average difference between its estimates and those according to the other two sources. Possible reasons for this—mostly related to differences in what each data source is aiming to estimate—are discussed further below.

Concurrence of County-level Estimates across Data Sources

With over 3,000 counties in the United States and thus limited sample size in many counties (ranging from 1 individual to 5,536 per county per month from the CTIS, with an average of 69.7 respondents), we selected the 460 most populous counties for

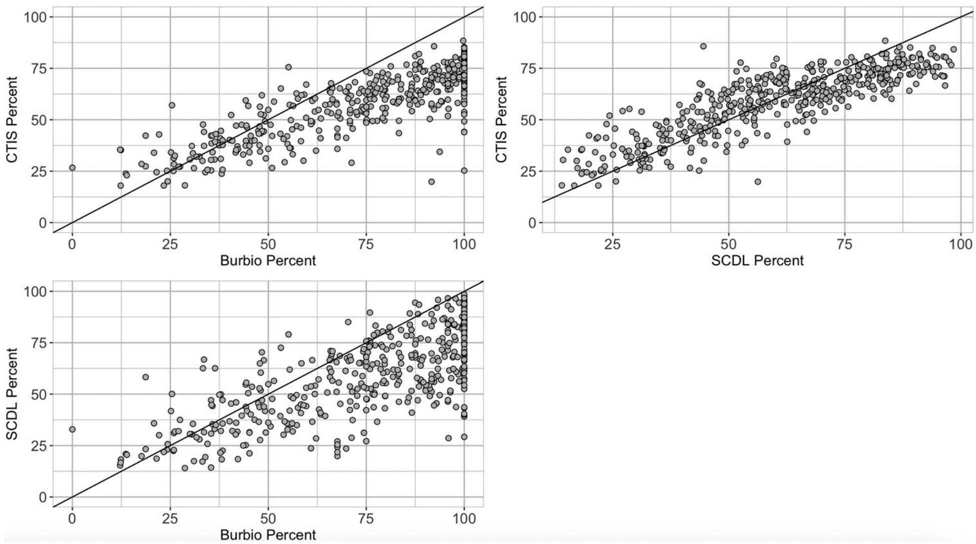

Figure 2. Percentage of students with any onsite instruction by county for CTIS, Burbio, and SCDL: largest 460 counties, averaged across entire time period.

Table 5. Correlations and RMSD between sources of county-level percentage of students with any onsite school.[a]

Any onsite school correlation (RMSD)	Burbio	CTIS	SCDL
Burbio	1 (0)	0.83 (28.13)	0.71 (29.78)
CTIS	0.83 (28.13)	1 (0)	0.82 (14.39)
SCDL	0.71 (29.78)	0.82 (14.39)	1 (0)

[a]Total sample size is n = 2,760 (460 counties * 6 months).

county-level comparisons. In this sample the sample sizes from the CTIS ranged from 30 to 5,536 per county per month, with an average of 329.6 respondents. Table 4 displays descriptive statistics for the largest 460 counties and all counties in the country. To visualize the consistency between the CTIS, Burbio, and SCDL percentages of any onsite instruction at the county level, Figure 2 shows three scatterplots. If full consistency is achieved, all points on the scatterplot would be on the y = x line, depicted in the plots. Each pairwise scatterplot shows strong grouping around this line, especially for the SCDL and CTIS scatterplot. Both scatterplots showing Burbio on the x-axis seem to indicate that Burbio tends to have higher estimates of onsite schooling percentages compared to CTIS and SCDL. This hypothesis is examined further below. To confirm the apparent relationships shown in the scatterplots, the percentage of students with any onsite school according to the CTIS was correlated at the monthly county level 0.83 ($p < .001$) with the Burbio measure and 0.82 ($p < .001$) with the SCDL measure, and the correlation between Burbio and SCDL percentages was 0.71 ($p < .001$) (Table 5). The RMSDs between the county-level estimates of onsite school are: 28.13 between CTIS and Burbio, 29.78 between SCDL and Burbio, and 14.39 between CTIS and SCDL. Similar to the state-level estimates, the correlations are relatively high but so are the RMSDs, and the CTIS and SCDL estimates have the closest values to one another on average.

Table 6. Results of multiple linear regression predicting the percentage of onsite school per county per month for all three county-level sources, weighted by county population.[a]

	Percentage of onsite		
Predictors	*Estimates*	*CI*	*p*
Intercept	44.53	42.30, 46.76	**<.001**
Source [Burbio]	4.16	1.59, 6.73	**.002**
Source [SCDL]	−0.26	−2.83, 2.31	.845
Month [Jan]	6.50	4.46, 8.53	**<.001**
Month [Feb]	13.49	10.97, 16.02	**<.001**
Month [Mar]	18.84	16.19, 21.49	**<.001**
Month [Apr]	22.65	19.99, 25.30	**<.001**
Month [May]	24.36	21.38, 27.34	**<.001**
% White	9.42	7.25, 11.59	**<.001**
% on public assistance	−3.26	−4.84, −1.67	**<.001**
% Working at home	−0.74	−2.45, 0.97	.395
GINI	1.73	0.09, 3.37	**.038**
Population density	0.59	−0.93, 2.10	.446
% Essential employees	3.69	1.93, 5.46	**<.001**
% with computers	−1.61	−3.69, 0.46	.127
% Adults with high school degree	−0.94	−3.58, 1.69	.482
Unemployment rate	−4.32	−5.68, −2.96	**<.001**
Cases per population	0.56	−0.48, 1.59	.291
Source [Burbio] * month [Jan]	−0.80	−3.67, 2.07	.583
Source [SCDL] * month [Jan]	7.47	4.60, 10.34	**<.001**
Source [Burbio] * month [Feb]	12.26	8.75, 15.77	**<.001**
Source [SCDL] * month [Feb]	0.37	−3.14, 3.88	.836
Source [Burbio] * month [Mar]	18.13	14.46, 21.79	**<.001**
Source [SCDL] * month [Mar]	3.52	−0.14, 7.18	.059
Source [Burbio] * month [Apr]	22.25	18.60, 25.89	**<.001**
Source [SCDL] * month [Apr]	1.11	−2.53, 4.76	.549
Source [Burbio] * month [May]	25.87	21.82, 29.91	**<.001**
Source [SCDL] * month [May]	−2.38	−6.43, 1.67	.249
Source [Burbio] * White	−1.20	−2.58, 0.18	.089
Source [SCDL] * White	0.46	−0.92, 1.85	.512
Source [Burbio] * % on public assistance	−5.21	−6.21, −4.20	**<.001**
Source [SCDL] * % on public assistance	−1.78	−2.78, −0.77	**.001**
Source [Burbio] * % working at home	2.45	1.39, 3.50	**<.001**
Source [SCDL] * % working at home	−0.57	−1.63, 0.48	.286
Source [Burbio] * GINI	−0.38	−1.39, 0.62	.452
Source [SCDL] * GINI	−3.18	−4.18, −2.17	**<.001**
Source [Burbio] * population density	0.73	−0.17, 1.63	.112
Source [SCDL] * population density	−1.13	−2.03, −0.23	**.014**
Source [Burbio] * % essential employees	1.67	0.58, 2.77	**.003**
Source [SCDL] * % essential employees	0.18	−0.92, 1.27	.750
Source [Burbio] * % with computers	−2.26	−3.55, −0.97	**.001**
Source [SCDL] * % with computers	−3.81	−5.10, −2.51	**<.001**
Source [Burbio] * % adults with high school degree	−2.09	−3.71, −0.47	**.012**
Source [SCDL] * % adults with high school degree	−3.35	−4.98, −1.73	**<.001**
Source [Burbio] * unemployment rate	−2.60	−3.86, −1.35	**<.001**
Source [SCDL] * unemployment rate	−3.99	−5.24, −2.73	**<.001**
Source [Burbio] * cases per population	2.63	1.24, 4.01	**<.001**
Source [SCDL] * cases per population	−1.62	−3.01, −0.24	**.022**
Random effects			
σ^2	3,134.57		
$\tau_{00\ \text{fips}}$	178.46		
ICC	0.05		
N_{fips}	460		
Observations	8,280		
Marginal R^2 / conditional R^2	0.111 / 0.159		

[a]Outcome of the model was percentage of any onsite school in each county and month (December, January, February, March, April, and May) for the three county-level sources (CTIS, Burbio, and SCDL). The model incorporated a factor indicating the source of the percentage, where CTIS was the reference category, a factor indicating month of data collection, as well as interaction terms between source and the other predictors in the model related to county-level characteristics and month. A random effect for county was also included. All continuous predictors were standardized, and the model was weighted by the logarithm of county population. Total sample size is n = 8,280 (460 counties * 6 months * 3 sources).
Bolded *p*-values indicate significance at the 0.05 level.

The average difference in the county-level percentages among the most populous 460 counties between CTIS and Burbio was −17.1 (95% CI: (−18.0, −16.3)), meaning that the CTIS reported percentage of households in the county with children attending school in person was, on average, 17.1 percentage points lower than the percentage of individuals in the county with schools offering onsite instruction according to Burbio. CTIS shows an average of 1.5% less onsite schooling per county per month compared to SCDL (95% CI: (−2.0, −1.0)), and Burbio reports an average of 15.6 percentage points more onsite schooling available per county per month than the SCDL's reported percentage of schools open in the county (95% CI: (14.7, 16.6)). Appendix Figure A2 displays these differences, broken down by pairwise comparison between sources and month. The overall average differences are also in Table A1 for both the county-level and state-level percentages.

Finally, we can discuss the results of the regression assessing whether county characteristics relate to the differences in county-level percentages of any onsite school across the CTIS, Burbio, and SCDL data (Table 6). We see that the average percentage of any onsite school offered in a county in a month as reported by Burbio was about 4.16% higher ($p = .002$) than the percentage of households with children attending onsite school reported by the CTIS in December (the reference month in the model), and the percentage of schools in the county open for onsite school according to SCDL was on average 0.26% lower ($p = .85$) than the percentage of households with children attending onsite school reported by the CTIS in December, holding all other variables at their mean and accounting for the county random intercept. For the CTIS percentages of households with children attending any onsite school and with all continuous variables at their mean, the average county monthly percentages of any onsite school increased each month, with January reporting 6.50% more onsite school than December ($p < .001$), all the way up to May reporting 24.36% more onsite school than December ($p < .001$). Several month/source interactions were statistically significant, primarily for the Burbio data, meaning that the pattern of differences between the Burbio and CTIS sources differed across time.

In terms of the county-level factors, for the CTIS data (the reference source), the percentage of households with children attending any onsite school in a county in a month was statistically significantly higher for counties with more White residents (9.42, $p < .001$), higher percentage of employees deemed essential (3.69, $p < .001$), and higher Gini coefficient (1.73, $p = .04$). On the other hand, the percentage of households with children attending any onsite school in a county in a month according to the CTIS was lower for counties with a higher unemployment rate (−4.32, $p < .001$) and more people on public assistance (−3.26, $p < .001$). There were several statistically significant interactions between source and month; for example, the coefficients of the interactions between Burbio and the months of February through May increased for each month, indicating that as time passed, the Burbio percentages became more different from the CTIS percentages. Other county-level factors also interacted with source, suggesting that the differences between reported onsite school percentages across sources varied somewhat due to factors at the county level like the percentage with computers, monthly COVID-19 cases per population, and county unemployment rate.

Discussion

There is substantial concern about the effects of online vs. onsite schooling on a number of factors, including COVID-19 transmission and academic, social, and emotional outcomes of children, their families, and their communities. In order to generate evidence regarding these important questions, we need consistent information about the extent of onsite schooling that has been occurring throughout the pandemic and how it varies across the country. This article provides the first documentation of the consistency across different measures of onsite schooling across the country, comparing primarily a large-scale Facebook platform-based survey of households, a county-level policy dataset, and mobile GPS data of school visits to assess similarities and differences.

The Delphi U.S. COVID-19 Trends and Impact Survey is unique in its national scope and household-level data. It reveals information to researchers about the decision-making of individual households and whether children were actually attending in-person school—whether it was public or private school. Similarly, the SCDL database informs about the foot traffic at schools, representing behaviors of individuals but aggregated to the school level. These data are a different method of measuring something similar to the CTIS data, so it is reassuring that we see similar results between the two. These sources could be used to answer similar questions, but the CTIS is more specific to individual children and households, while the SCDL is geared toward schools. Also importantly, the SCDL data include only public schools and include data from staff and families.

The Burbio data have a different purpose and level of measurement, in that it is measuring school policy rather than individual behavior. Given that even when schools were "open" for in-person schooling, many areas allowed families to opt-out and keep their children home, it is thus not surprising that the Burbio data generally report a higher percentage of onsite school. This is one explanation for why the correlations between all three sources are similar—indicating similar patterns in information—but with the Burbio data showing higher RMSD's when compared to the other two sources. Burbio percentages were generally higher due to the difference in the measurement just described and discussed further below. Like the SCDL, the Burbio data are only available for public schools.

These findings indicate broad consistency between the datasets shown in the high state- and county-level correlations, but also show some differences between data sources, which is expected on some level due to the differences in units of measurement (i.e., households, schools, districts, and counties) and data collection goals. From Appendix Figure A2, we see that the average difference between CTIS and SCDL is very close to 0 for all months, and the data are somewhat normally distributed around 0. These sources are therefore somewhat similar, which aligns with expectations since CTIS captures household-level behavior and SCDL captures individual-level behavior aggregated to the school level. Again, by nature of Burbio's data collection, it makes sense for the percentages to be higher than the other two sources, because Burbio reports on the county policies and the onsite schooling available to students rather than the individual-level behavior of households.

On the other hand, county measures that instead have higher percentages of onsite school according to CTIS compared to Burbio are primarily from the months of December, January, and February, and are often for the counties that Burbio denotes as fully online or mostly online. Some of the variation here could be due to private/Catholic schools in the CTIS data that might have higher percentages of onsite school compared to solely public schools that are reported on by Burbio. Approximately 10.2% of all children in elementary or secondary schools are attending private schools (NCES, 2019), so the sample from Facebook likely includes some families sending their children to private school. Although private versus public school is not directly addressed in the CTIS, there are some differences between the CTIS sample and the national population. Specifically, in comparing the sample from the CTIS with the 2019 American Community Survey sample, the CTIS sample overrepresents higher education levels (Salomon et al., 2021). These results highlight the distinctions between the levels of data collection of each dataset and show why researchers should consider their research questions and unit of analysis when selecting which dataset might be most helpful to utilize.

Some associations were found when examining the role of county-level factors in both levels of in-person schooling and the differences across sources. In particular, the percentage of any onsite school increased over time. Furthermore, race, public assistance, Gini coefficient, essential employees, and unemployment rate were associated with onsite schooling percentages. Many interaction terms were also statistically significant specifically for Burbio, indicating that the difference in households with children attending school onsite reported by CTIS compared to overall availability of onsite school according to county policy reported by Burbio varied across some county factors such as computer access, education levels in the county, and unemployment rate. These county-level factors can be further explored to determine more about how they relate to individual families' decisions to send their children to school or not. For example, higher percentages of the county receiving public assistance is associated with lower percentage of onsite schooling for all three sources, and this negative relationship is strongest in the Burbio dataset and weakest in the CTIS. Part of this interaction effect could be related to private schooling, which is included in the CTIS but not in Burbio or SCDL.

While this study is the first that assesses consistency between data sources and helps motivate future research on the effects of school policies during the COVID-19 pandemic, a limitation of the study is that it gives only a snapshot of the levels of consistency, examining winter of 2020 and spring of 2021. Other patterns may have been seen in early Fall 2020, but there was less variation at this time across regions, and a large portion of the data used in this analysis were not sufficiently available at that time. This work can extend into the 2021–2022 school year as well, but we ultimately decided against including that data because the percentages of onsite school in the country moved much closer to 100%. Specifically, during the time frame of our study (December 2020–May 2021), the average percentage of households in the United States with children attending classes full-time according to the CTIS (a slightly different question than the one considered in the majority of this article which also looked at part-time school) was 42.1%. In comparison, this percentage for September 2021 to December 2021 was 85.1%, and for February 2022 to May 2022 was 94.5%.

An additional limitation is that the CTIS samples only Facebook users and might not be fully representative of the general population. The CTIS also does not distinguish the type of schooling (e.g., public, private, charter), whereas Burbio and SCDL focus on public schools—this could be a strength or limitation depending on the research question. Due to sample size limitations, the county-level analyses focused on the largest counties in the country; implications for small counties are unclear. We conjecture that the overall consistency across data sources would be similar in these smaller counties, but with substantial variability and uncertainty making the comparisons more challenging. We can therefore only truly apply our conclusions to the large counties discussed in this article. This article also did not use all datasets available in this field of school policies, such as the COVID-19 School Data Hub (Oster et al., 2021) because we could not efficiently and clearly compare all datasets at the same level of regional and time aggregation. Future work can investigate some of these datasets as well and see how their estimates compare. An important final limitation is that without a gold-standard source indicating the true state of schooling experiences, showing consistency of data sources that do exist does not necessarily mean that all of the data sources are completely accurate. However, the consistency across multiple sources does provide some reassurance that the sources are capturing a similar view that is hopefully close to the truth.

Most crucially, the sort of examination and comparison here serves to clarify what each source is measuring, thus helping researchers select the best source for their research question, with some confidence that the sources are measuring the same underlying constructs. For example, when studying household-level mental health and its relation to onsite schooling, it is important to use information on household-specific schooling modalities. In contrast, when studying the effects of school modality policies on COVID-19 or academic outcomes more broadly, the Burbio data, which reflects actual school policies, may be most appropriate. In other analyses, any of the sources could be useful, or a researcher could select one source but use the others to assess consistency of findings.

It is important to make a choice to use data depending on the goal level of analysis (e.g., household versus county) because different datasets could lead to slightly different conclusions. Therefore, the differences between datasets could matter in practice, highlighting the continuing need for more research and more data collection on schooling experiences during the COVID-19 pandemic. The broad consistency seen across sources in this work does provide reassurance that current measures of onsite schooling can be utilized to help learn about schooling modalities during the pandemic. Moving forward with this knowledge, researchers can begin to use existing data to investigate questions related to the impact of educational policies on children's learning and development.

Open Research Statements

Study and Analysis Plan Registration

- There is no study and analysis plan registration associated with this manuscript.

Data, Code, and Materials Transparency

- Some of the data used in this study are available publicly on the Open Science Framework (https://osf.io/tpwqf/), the U.S. Census Bureau (https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-detail.html), and MCH (https://www.mchdata.com/covid19/school-closings). Remaining datasets can be requested from the Delphi Group at CMU (https://delphi.cmu.edu/covidcast/) and Burbio (https://about.burbio.com/school-opening-tracker). The code used to derive the findings of this study are available in a public repository on GitHub (https://github.com/carlyls/covid19_schooldata/).

Design and Analysis Reporting Guidelines

- This manuscript was not required to disclose use of reporting guidelines, as it was initially submitted prior to JREE mandating open research statements in April 2022.

Transparency Declaration

- The lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Replication Statement

- This manuscript reports an original study.

Acknowledgements

This research is partially based on survey results from Carnegie Mellon University's Delphi Group. We also thank Burbio, the U.S. School Closure and Distance Learning Database, and MCH for data access.

Funding

This work was supported by Johns Hopkins Epidemiology and Biostatistics of Aging T32 NRSA Training Grant; Johns Hopkins University Discovery Award;Johns Hopkins University COVID-19 Modeling and Policy Hub Award.

ORCID

Carly Lupton-Smith http://orcid.org/0000-0003-1067-3260

References

AEI. (2022, June 3). *Enrollment tracker: 2020–2022. Return to learn tracker*. https://www.returntolearntracker.net/

Armitage, R., & Nellums, L. B. (2020). Considering inequalities in the school closure response to COVID-19. *The Lancet: Global Health*, *8*(5), e644. https://doi.org/10.1016/S2214-109X(20)30116-9

Auger, K. A., Shah, S. S., Richardson, T., Hartley, D., Hall, M., Warniment, A., Timmons, K., Bosse, D., Ferris, S. A., Brady, P. W., Schondelmeyer, A. C., & Thomson, J. E. (2020). Association between statewide school closure and COVID-19 incidence and mortality in the US. *JAMA*, *324*(9), 859. https://doi.org/10.1001/jama.2020.14348

Bansak, C., & Starr, M. (2021). Covid-19 shocks to education supply: How 200,000 U.S. households dealt with the sudden shift to distance learning. *Review of Economics of the Household*, *19*(1), 63–90. https://doi.org/10.1007/s11150-020-09540-9

Barkay, N., Cobb, C., & Eilat, R. (2020). *Weights and methodology brief for the COVID-19 Symptom Survey by University of Maryland and Carnegie Mellon University, in partnership with Facebook*. Advance online publication. arXiv. https://doi.org/10.48550/arXiv.2009.14675

Berger, F., Schreiner, C., Hagleitner, W., Jesacher-Rößler, L., Roßnagl, S., & Kraler, C. (2021). Predicting coping with self-regulated distance learning in times of COVID-19: Evidence from a longitudinal study. *Frontiers in Psychology*, *12*, 701255. https://doi.org/10.3389/fpsyg.2021.701255

Burbio. (2020a). *Burbio's K-12 school opening tracker*. https://about.burbio.com/school-opening-tracker

Burbio. (2020b). *Burbio's methodology*. https://about.burbio.com/methodology/

Center for Disease Control and Prevention. (2021). *Operational strategy for K-12 schools through phased mitigation*. https://www.cdc.gov/coronavirus/2019-ncov/community/schools-childcare/operation-strategy.html#

Center for Reinventing Public Education. (2020). *Fall 2020: The state of school reopening*. https://www.crpe.org/current-research/covid-19-school-closures.

DATA: State Dashboards on COVID-19 in Schools and Instructional Models. (2020, December 30). https://www.edweek.org/policy-politics/data-state-dashboards-on-covid-19-in-schools-and-instructional-models/2020/11

Delphi Group. (2020). *COVID symptom survey*. https://cmu-delphi.github.io/delphi-epidata/symptom-survey/

Delphi Group. (2022, March 1). *COVID-19 trends and impact survey (CTIS) RESULTS*. Carnegie Mellon University Delphi Group. https://delphi.cmu.edu/covidcast/survey-results/

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *118*(17), e2022376118. https://doi.org/10.1073/pnas.2022376118

Eyles, A., Gibbons, S., & Montebruno, P. (2020). Covid-19 school shutdowns: What will they do to our children's education? *CEP COVID-19 Analysis*. https://cep.lse.ac.uk/pubs/download/cep-covid-19-001.pdf

Falk, A., Benda, A., Falk, P., Steffen, S., Wallace, Z., & Hoeg, T. B. (2021). COVID-19 cases and transmission in 17 K–12 schools—Wood County, Wisconsin, August 31–November 29, 2020. *MMWR. Morbidity and Mortality Weekly Report*, *70*(4), 136–140. https://doi.org/10.15585/mmwr.mm7004e3

Glenn, E. H. (2019). *Package 'acs'*. https://cran.r-project.org/web/packages/acs/acs.pdf

Goldhaber, D., Imberman, S., Strunk, K., Hopkins, B., Brown, N., Harbatkin, E., & Kilbride, T. (2020). *To what extent does in-person schooling contribute to the spread of COVID-19? Evidence from Michigan and Washington*. Education Policy Innovation Collaborative..

Hamilton, L. S., & Ercikan, K. (2022). COVID-19 and US schools: Using data to understand and mitigate inequities in instruction and learning. In F. M. Reimers (Ed.), *Primary and secondary education during Covid-19* (pp. 327–351). Springer, Cham.

Harris, D. N., Ziedan, E., & Hassig, S. (2021). *The effects of school reopenings on COVID-19 hospitalizations*. National Center for Research on Education Access and Choice.

Hartney, M. T., & Finger, L. K. (2022). *Politics, markets, and pandemics: public education's response to COVID-19. Perspectives on Politics, 20*(2), 457–473. https://doi.org/10.1017/S1537592721000955

Honein, M. A., Barrios, L. C., & Brooks, J. T. (2021). Data and policy to guide opening schools safely to limit the spread of SARS-CoV-2 infection. *JAMA*, *325*(9), 823. https://doi.org/10.1001/jama.2021.0374

Institute of Education Sciences. (2022, April). *2022 School pulse panel*. https://ies.ed.gov/school-survey/spp/#tab-1

Jehn, M., McCullough, J. M., Dale, A. P., Gue, M., Eller, B., Cullen, T., & Scott, S. E. (2021). Association between K-12 school mask policies and school-associated COVID-19 outbreaks - Maricopa and pima counties, Arizona, July-August 2021. *MMWR: Morbidity and Mortality Weekly Report*, *70*(39), 1372–1373. https://doi.org/10.15585/mmwr.mm7039e1

Johns Hopkins Berman Institute of Bioethics. (2020, June 29). *School reopening plan tracker*. https://bioethics.jhu.edu/research-and-outreach/projects/eschool-initiative/school-policy-tracker/

Kreuter, F. (2020). Partnering with Facebook on a university-based rapid turn-around global survey. *Survey Research Methods*, *14*(2), 159–163. https://doi.org/10.18148/srm/2020.v14i2.7761

Lessler, J., Grabowski, M. K., Grantz, K. H., Badillo-Goicoechea, E., Metcalf, C., Lupton-Smith, C., Azman, A. S., & Stuart, E. A. (2021). Household COVID-19 risk and in-person schooling. *Science*, *372*(6546), 1092–1097. https://doi.org/10.1126/science.abh2939

Martin, A. J., Collie, R. J., & Nagy, R. P. (2021). Adaptability and high school students' online learning during COVID-19: A job demands-resources perspective. *Frontiers in Psychology*, *12*, 702163. https://doi.org/10.3389/fpsyg.2021.702163

MCH Strategic Data. (2020). *COVID-19 impact: School district status updates for spring 2021*. https://www.mchdata.com/covid19/schoolclosings

McLoughlin, G. M., McCarthy, J. A., McGuirt, J. T., Singleton, C. R., Dunn, C. G., & Gadhoke, P. (2020). Addressing food insecurity through a health equity lens: A case study of large urban school districts during the COVID-19 pandemic. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, *97*(6), 759–775. https://doi.org/10.1007/s11524-020-00476-0

National Center for Education Statistics. (2019). *School choice in the United States: 2019*. https://nces.ed.gov/programs/schoolchoice/ind_03.asp

Oster, E., Halloran, C., & Jack, R. (2021). COVID-19 school data hub. https://www.covidschooldatahub.com/

Oster, E., Jack, R., Halloran, C., Schoof, J., McLeod, D., Yang, H., Roche, J., & Roche, D. (2021). Disparities in learning mode access among K–12 students during the COVID-19 pandemic, by race/ethnicity, geography, and grade level—United States, September 2020–April 2021. *MMWR. Morbidity and Mortality Weekly Report*, *70*(26), 953–958. https://doi.org/10.15585/mmwr.mm7026e2

Parolin, Z., & Lee, E. (2021). Large socio-economic, geographic and demographic disparities exist in exposure to school closures. *Nature Human Behaviour*, *5*(4), 522–528. https://doi.org/10.1038/s41562-021-01087-8

Parolin, Z., & Lee, E. (2022, January 10). *U.S. school closure & distance learning database*. https://doi.org/10.17605/OSF.IO/TPWQF

Salomon, J. A., Reinhart, A., & Bilinski, A. (2021). The US COVID-19 Trends and Impact Survey: Continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination. *Proceedings of the National Academy of Sciences of the United States of America*. *118*(51), e2111454118. https://doi.org/10.1073/pnas.2111454118

Sitthiyot, T., & Holasut, K. (2020). A simple method for measuring inequality. *Palgrave Communications*, *6*(1), 1–9. https://doi.org/10.1057/s41599-020-0484-6

University of Southern California. (2020). *Understanding America study*. https://covid19pulse.usc.edu/

USA Facts. (2021, January 30). *US Coronavirus cases and deaths*. https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/

U.S. Bureau of Labor Statistics. (2021). *Overview of BLS statistics on unemployment*. https://www.bls.gov/bls/unemployment.htm

U.S. Census Bureau. (2020a, June 22). *County population by characteristics: 2010-2019*. https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-detail.html

U.S. Census Bureau. (2020b, December 07). *Urban and rural*. https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html

U.S. Census Bureau. (2021, January 12). *Measuring household experiences during the Coronavirus pandemic*. https://www.census.gov/data/experimental-data-products/household-pulse-survey.html

USDA ERS. (2021, January 5). *Download data*. https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/

Van Lancker, W., & Parolin, Z. (2020). COVID-19, school closures, and child poverty: A social crisis in the making. *The Lancet Public Health*, *5*(5), e243–e244. https://doi.org/10.1016/S2468-2667(20)30084-0

World Population Review. (2021). *US states - Ranked by population 2021*. https://worldpopulationreview.com/states

Ykzeng. (2020, April 20). *Ykzeng/covid-19*. https://github.com/ykzeng/covid-19/tree/master/data

Zhang, K., Wu, S., Xu, Y., Cao, W., Goetz, T., & Parks-Stamm, E. J. (2020). Adaptability promotes student engagement under COVID-19: The multiple mediating effects of academic emotion. *Frontiers in Psychology*, *11*, 633265. https://doi.org/10.3389/fpsyg.2020.633265

Zimmerman, K. O., Akinboyo, I. C., Brookhart, M. A., Boutzoukas, A. E., McGann, K. A., Smith, M. J., Maradiaga Panayotti, G., Armstrong, S. C., Bristow, H., Parker, D., Zadrozny, S., Weber, D. J., & Benjamin, D. K. (2021). Incidence and secondary transmission of SARS-CoV-2 infections in schools. *Pediatrics*, *147*(4), e2020048090. https://doi.org/10.1542/peds.2020-048090

Appendix A

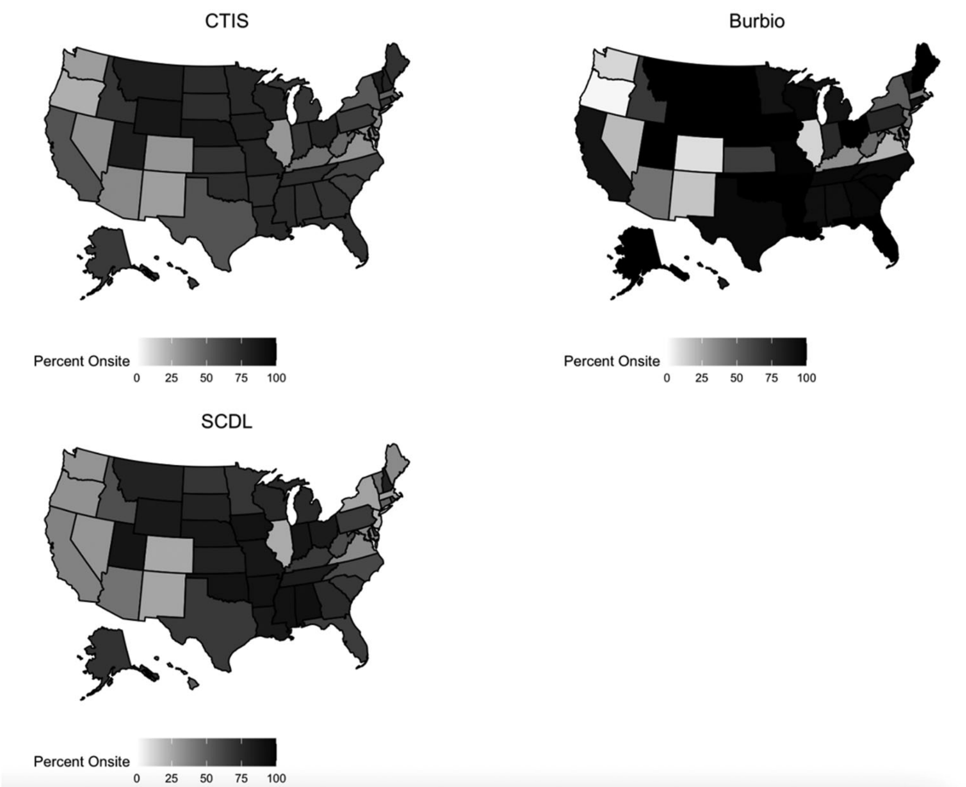

Figure A1. Average percentage of onsite instruction by state for each source. The precise definitions of the percentages according to each source are: For CTIS, the percentage of households with children attending any onsite school per state; for Burbio and SCDL, the percentage of students in the state attending a school that is offering any onsite instruction.

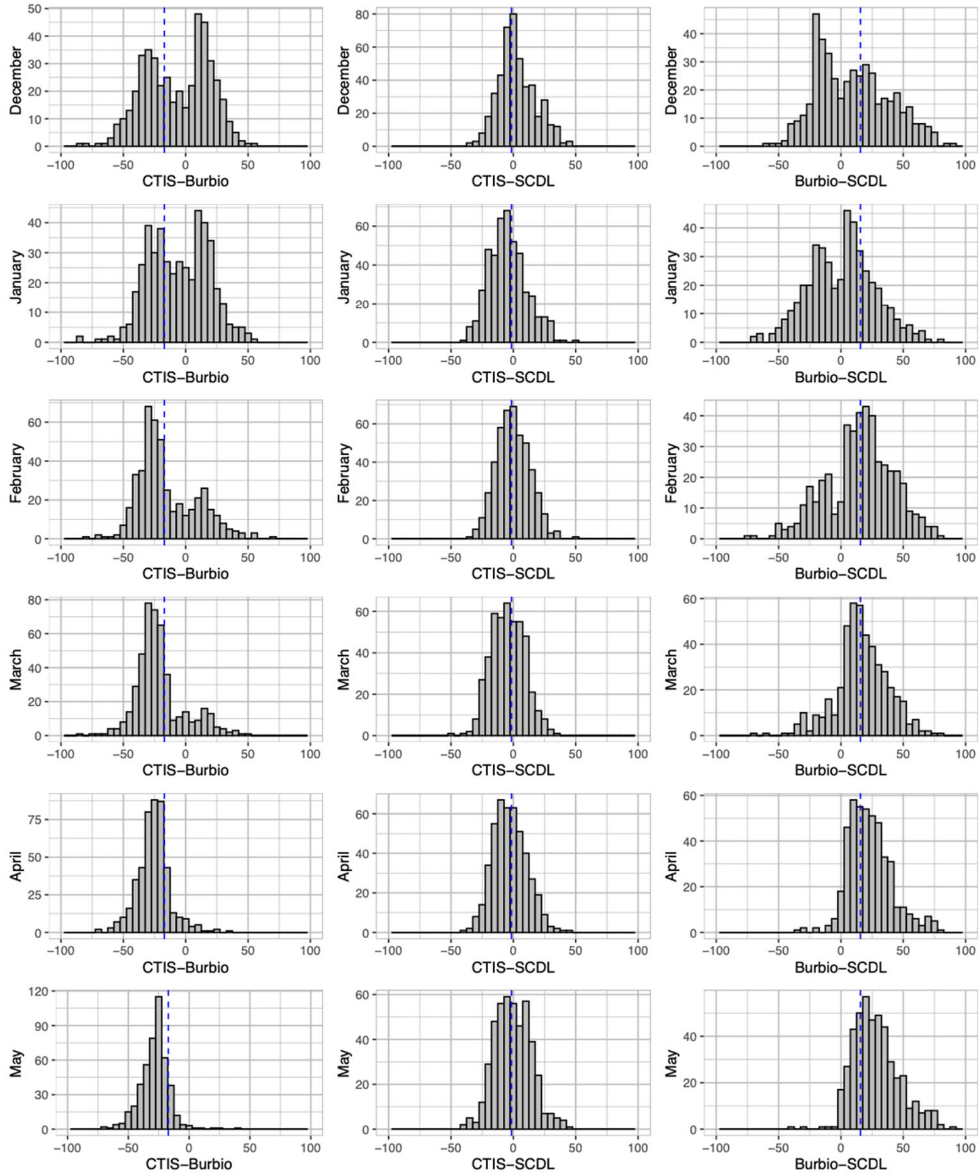

Figure A2. Monthly differences between sources for measures of percentage of students with any onsite school by county, largest 460 counties.

Table A1. Descriptive statistics for the difference in reported onsite school percentage for all states and the largest 460 counties.

	CTIS-Burbio mean (SD)	CTIS-SCDL mean (SD)	Burbio-SCDL mean (SD)
States	−13.6 (16.5)	−1.8 (11.1)	11.8 (18.4)
Largest 460 counties	−17.1 (22.3)	−1.5 (14.3)	15.6 (25.4)

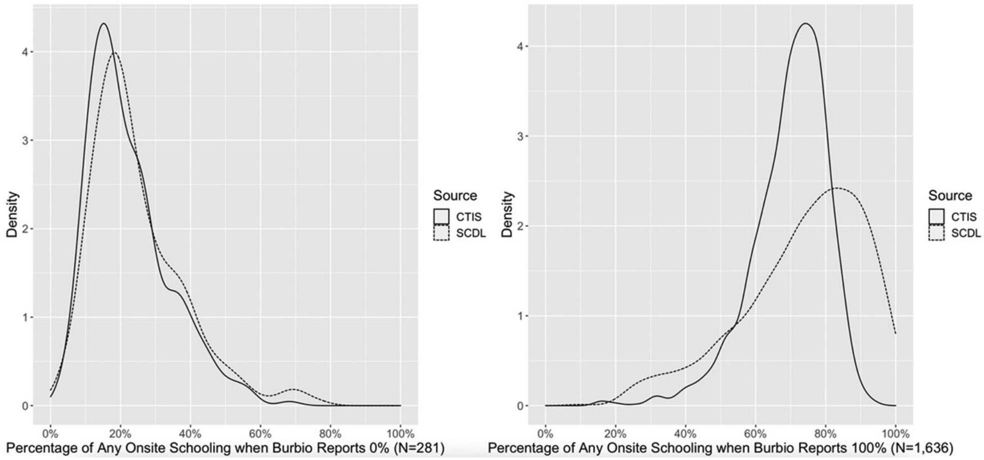

Figure A3. The distribution of percentage of any onsite schooling for counties that had 0% or 100% onsite school in a month according to Burbio.

Table A2. Descriptive statistics of districts classified by MCH.[a]

MCH classification	Number of districts	Burbio mean % onsite (SD)	SCDL mean % onsite (SD)
On premises	16	87 (29.6)	51.9 (25.1)
Hybrid	75	76 (33.2)	58.6 (28.7)
Online only	73	12.8 (26.5)	29.1 (21.9)
Other	38	68.7 (40.8)	55.4 (26.3)

[a]Total sample size is n = 202 districts.