IMPORTANCE The consequences of school closures for children's health are profound, but existing evidence on their effectiveness in limiting severe acute respiratory syndrome coronavirus 2 transmission is unsettled.

OBJECTIVE To determine the independent associations of voluntary behavioral change, school closures, and bans on large gatherings with the incidence and mortality due to coronavirus disease 2019 (COVID-19).

DESIGN, SETTING, AND PARTICIPANTS This population-based, interrupted-time-series analysis of lagged independent variables used publicly available observational data from US states during a 60-day period from March 8 to May 18, 2020. The behavioral measures were collected from anonymized cell phone or internet data for individuals in the US and compared with a baseline of January 3 to February 6, 2020. Estimates were also controlled for several state-level characteristics.

EXPOSURES Days since school closure, days since a ban on gatherings of 10 or more people, and days since residents voluntarily conducted a 15% or more decline in time spent at work via Google Mobility data.

MAIN OUTCOMES AND MEASURES The natural log of 7-day mean COVID-19 incidence and mortality.

RESULTS During the study period, the rate of restaurant dining declined from 1 year earlier by a mean (SD) of 98.3% (5.2%) during the study period. Time at work declined by a mean (SD) of 40.0% (7.9%); time at home increased by a mean (SD) of 15.4% (3.7%). In fully adjusted models, an advance of 1 day in implementing mandatory school closures was associated with a 3.5% reduction (incidence rate ratio [IRR], 0.965; 95% CI, 0.946-0.984) in incidence, whereas each day earlier that behavioral change occurred was associated with a 9.3% reduction (IRR, 0.907; 95% CI, 0.890-0.925) in incidence. For mortality, each day earlier that school closures occurred was associated with a subsequent 3.8% reduction (IRR, 0.962; 95% CI, 0.926-0.998), and each day of advance in behavioral change was associated with a 9.8% reduction (IRR, 0.902; 95% CI, 0.869-0.936). Simulations suggest that a 2-week delay in school closures alone would have been associated with an additional 23 000 (95% CI, 2000-62 000) deaths, whereas a 2-week delay in voluntary behavioral change with school closures remaining the same would have been associated with an additional 140 000 (95% CI, 65 000-294 000) deaths.

CONCLUSIONS AND RELEVANCE In light of the harm to children of closing schools, these findings suggest that policy makers should consider better leveraging the public's willingness to protect itself through voluntary behavioral change.
Few measures taken to combat the coronavirus disease 2019 (COVID-19) pandemic are more far-reaching and consequential than closing schools, which disrupts children's socioemotional learning and academic performance as well as parents' employment and education. The economic costs of school closures in the US in the spring of 2020 have been estimated to be some $2 trillion dollars and fall most heavily on those least able to withstand them. However, although schools are sites of influenza transmission, the evidence on school closures and reduced COVID-19 spread is inconclusive.

Several peer-reviewed studies find that the independent association of school closures with subsequent rates of COVID-19 is low to nonexistent. By contrast, a recent study by Auger and coauthors finds large associations of school closures with lower COVID-19 incidence and mortality. Using the same data and similar methods, Courtemanche and coauthors come to the opposite conclusion: school closures are associated with an increase in COVID-19 cases and mortality, albeit not significantly. Neither of these articles control for the contemporaneous effects of voluntary behavioral change. It may be that fear of contagion was an underlying factor for both individual behavioral change and public policy change.

This study tests the association of school closures with subsequent changes in COVID-19 incidence and deaths while controlling for contemporaneous voluntary behavioral change. The analysis adds further evidence to the policy debate by incorporating big data to analyze the timing of government policies such as school closures, closures of nonessential businesses, and stay-at-home orders relative to several measures of behavioral change in the general population. The analysis uses data-driven techniques to incorporate these co-occurring policy and behavioral changes while minimizing the risk for collider bias. It also controls for serial autocorrelation in the estimates.

Methods

For this cross-sectional study, an interrupted-time-series model was fit to state-level data on COVID-19 incidence and deaths and on several legal restrictions and measures of behavioral change. All data were deidentified. The study was deemed exempt from institutional review by the University of California, Los Angeles institutional review board. Informed consent was not required for the use of publicly available data. This analysis follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.

Outcome Measures

Data on COVID-19 cases and deaths were taken from the New York Times. Population data from the 2018 American Community Survey were used to calculate rates per 1000 population. The primary outcome is the logged 7-day rolling mean of daily incidence and death rate, with values of 0 recoded to 0.001.

Primary Regressors of Interest

Measures for policy action take the form of the number of days since the action was implemented (eTable 1 in the Supplement). Candidate policy variables included school closure, nonessential business closures, bans on large public gatherings, stay-at-home orders, and restaurant closures.

Several proxies for behavioral change were collected, coded as the number of days since the behavioral change occurred, which was assessed as crossing a percentage of change threshold from baseline values. Time at work and time at home were assessed from Google Mobility, gathered from individual cell phone data. These variables were coded as the percentage of change relative to a pre–COVID-19 baseline of median values from January 3 to February 6, 2020. Personal hygiene behaviors were the most difficult to observe using big data, so the number of daily Google searches, by state, for the term hand sanitizer served as a proxy. Google search data have been previously used in epidemiologic studies, with good predictive validity for observed behavior at the state level. This term was chosen because of the consistency of public health advice (compared with mask use, for example) and because the use of hand sanitizer would require most people to obtain a product (unlike hand washing, for example). State values were normalized by dividing the number of searches by the count on the date with the most frequent searches for that term. Daily restaurant dining data were taken from OpenTable. Data were coded as the percentage of change in diners in 2020 relative to the same date in 2019. OpenTable data were available only for the 36 most populous states. Because daily data tend to have high variance, all variables were converted to 7-day moving means.

Covariates

Consistent with previous work, regression analyses controlled for a set of state-level covariates:

• Percentages of the population younger than 15 years and older than 65 years;
• Percentage of the population in nursing homes;
• Quartile of cumulative incidence growth rate during the first 10 days of the pandemic;
• The 2018 Centers for Disease Control and Prevention Social Vulnerability Index, reweighted at the state level;
• Cumulative testing rates of less than 10, 10 to 20, and greater than 20 per 1000 population;
• Obesity rate.
• Urban density, measured by the number of urban residents per square mile (<50, 50-100, 100-150, and >150)\textsuperscript{2}; and
• Number of days into the study period. To adjust for the fact that the pandemic hit different states on different weeks, we recentered our data by defining all time-varying data according to the number of days since the pandemic began, defined as the first day on which a state had a cumulative incidence rate of 0.5 cases per 100,000 population or higher. The study period accordingly varied by state in calendar time and was defined as the 60-day period after the beginning date.

**Statistical Analysis**

To understand the sequencing of behavioral change and policies, each of the 4 behavioral measures was graphed against time. To help with interpretability, this graph is centered on days since school closures went into effect rather than days into the study period. Also indicated on the graphs are the dates (relative to school closures) of legal restrictions most closely related to the behavioral measure.

To limit problems associated with collider bias among the legal restrictions in the regression analysis,\textsuperscript{26} we tested the association of cumulative COVID-19 incidence at the end of the study period with indicators of legal restrictions at the state level. State-level rates increased rapidly in the early stages and then abruptly leveled off around day 30, so if this pattern did not vary according to whether a state implemented a given legal restriction, then that policy was assessed as not a plausible confounder and therefore excluded from further analyses to avoid collider bias.

Of the 5 legal restrictions identified, 3 (nonessential business closure, restaurant closure, and stay-at-home orders) were implemented in some states but not others. Because none of the restrictions with variation across states were significant at the 0.05 level in the regressions (eTable 2 in the Supplement), they were dropped from subsequent analysis. School closure and bans on large gatherings were retained as candidates, before these other interventions took effect. Restaurant dining, for example, was reduced 50% to 100% before any restaurant closures took effect.

Because of differences in testing capacity and baseline risk across states, we used a data-driven process with the bayesian information criteria to determine lag periods, selection of the behavior-change variable to include, and candidate thresholds for the behavioral change variable (eTable 5 in the Supplement). Based on the results across the 2 outcomes, we selected 15% decrease in time spent at work. Lag periods were not forced to be identical across the policy and behavior variables, but the highest-performing models all had nearly identical lags across the measures: 14 days for the incidence models and 21 days for the death models.

To adjust for serial autocorrelation in the dependent variable and therefore in the error terms, a second set of models were estimated using the Prais-Winsten estimator.\textsuperscript{27} All analyses were conducted in STATA, version 15 (StataCorp LLC).

**Results**

There was wide variation in the behaviors during the 60-day periods in each state ranging from March 8 to May 7 to March 19 to May 18, 2020, with the rate of meals in restaurants declining by a mean (SD) of 98.3% (5.2%) during the study period and a maximum of 100% since the previous year; time at work declining by a mean (SD) of 40.0% (7.9%) and a maximum of 60% since the pre-COVID baseline; and time at home increasing by a mean (SD) of 15.4% (3.7%) and a maximum of 26% since the pre-COVID baseline. eTable 6 in the Supplement reports the descriptive statistics of the COVID-19-related behaviors during the study period.

Figure 1 shows the patterns of these changes against the timing of school closures and other legal restrictions. The experience of each state is represented separately by its own line. Behavioral changes clearly began before school closure across the 50 states. In many states, the behavioral change was about halfway to its full extent by the time schools were officially closed. Because the other policy changes, such as the closure of nonessential businesses and stay-at-home orders, occurred alongside or after school closures, the behavioral changes were well under way, and in some cases fully realized, before these other interventions took effect. Restaurant dining, for example, was reduced 50% to 100% before any restaurant closures took effect.

There was large variation in the spread of the pandemic during the first 10 days: median growth rate was just over 1300% and ranged from 180% in South Dakota to more than 5000% in New Jersey. Behavioral change began as early as 4 days before the pandemic was under way (as defined by the study period) and as late as 10 days afterward. eTable 7 in the Supplement displays characteristics of the states during the study period. Other state-level characteristics are presented in eTable 8 in the Supplement.

Figure 2 and Figure 3 show 7-day rolling mean daily incidence and mortality by state. States are grouped by tertile based on their cumulative incidence at the time of the policy or behavioral change. The dashed blue lines indicate at which day of the pandemic the respective policy or behavioral change occurred. For all 3 sets of actions, states that enacted policies or
changed behaviors early in the pandemic generally had lower incidence and death throughout the study period.

The Table shows the results for the 8 regression models. The upper half of the table displays the models using logged daily incidence as the outcome, with the Prais-Winsten models as our preferred estimation method. In model 1, in which only school-closure timing and state-level controls are included, each day earlier that schools were closed was associated with an incidence rate ratio (IRR) of 0.854 (95% CI, 0.845-0.863). This finding implies that each day earlier that schools were closed was associated with a 14.6% reduction in cases. In model 2, which adds the number of days since the ban on gatherings, the association between earlier school closures and cases weakened by more than one-fourth (IRR, 0.892; 95% CI, 0.880-0.905). The IRR for gathering bans was smaller but still significant (0.949; 95% CI, 0.940-0.960). Model 3 removed both policy variables and introduced the behavior-change variable, finding that each additional day since a 15% reduction in time spent at work was associated with a 15.3% (IRR, 0.847; 95% CI, 0.838-0.856) reduction in daily COVID-19 incidence. Model 4 is the preferred specification, because it included all 3 primary regressors. Each day of delay in school closures was associated with a 3.5% reduction (IRR, 0.965; 95% CI, 0.946-0.984), whereas each day earlier that behavioral change occurred was associated with a 9.3% reduction (IRR, 0.907; 95% CI, 0.890-0.925).

The bottom half of the Table shows the results for the daily death rate. The findings were similar to those for daily incidence across the 4 Prais-Winsten models: school closures 1 day earlier were associated with a 15.3% reduction (IRR, 0.847; 95% CI, 0.829-0.865) in daily death in the short model, which fell to a 3.8% reduction (IRR, 0.962; 95% CI, 0.926-0.998) when the other primary regressors were accounted for. For the reduction in time spent at work, these estimates were a 15.9% reduction (IRR, 0.841; 95% CI, 0.824-0.860) in the short model and a 9.8% reduction (IRR, 0.902; 95% CI, 0.869-0.936) in daily deaths in the full model. Figure 4 shows the results of simulations under various scenarios of daily incidence and cumulative cases and for deaths. The baseline scenario was the estimated value under the full model, which was close to the actual levels of incidence and death (eFigure in the Supplement). In each graph, 3 hypothetical counterfactuals were graphed:

---

**Figure 1. Timing of Coronavirus Disease 2019-Related Behavioral Changes in Association With School Closures**

A total of 3000 state-day values for each variable as defined in the text are shown, with 1 line connecting points in each state, except for dining data, which were not available for Alaska, Arkansas, Delaware, Idaho, Iowa, Maine, Mississippi, Montana, New Hampshire, North and South Dakota, Vermont, West Virginia, and Wyoming.
• The modeled values if there had been a 2-week delay in school closures in each state relative to the actual closure date but with behavioral change happening when it did;
• The modeled values if there had been a 2-week delay in behavioral change in each state relative to the actual timing of behavioral change but with school closures happening when they did; and
• The modeled values if both school closures and behavioral change had happened with a 2-week delay.

These graphs suggest that had the school closures happened with a 2-week delay, there would have been approximately 587,000 (95% CI, 213,000-1,153,000) more cases and 23,000 (95% CI, 20,000-62,000) more deaths by day 60 of the pandemic than in fact occurred. Had the behavioral changes happened with a 2-week delay, there would have been some 4.3 million (95% CI, 2,542,000-6,995,000) more cases and 140,000 (95% CI, 65,000-294,000) more deaths than in fact occurred. Had both changes happened 2 weeks later than they did, there would have been 10.1 million (95% CI, 4,841,000-19,944,000) more cases and 280,000 (95% CI, 105,000-788,000) more deaths.

Discussion
This analysis shows that a substantial portion of behavioral changes that were eventually embraced happened before any...
intervention requiring such action. The results imply that many people will act in ways that limit the spread of a pandemic disease even without being compelled to do so.

The regression analyses suggest that the school closures of mid-to-late March 2020 appear to have been associated with a reduced incidence of COVID-19 infection and associated deaths due to COVID-19. However, relative to other changes occurring in the same time frame, the estimates of association were modest. The simulations suggested that timely school closures were associated with the prevention of cases and deaths during the study period, but that the estimated number of cases and deaths averted because of voluntary changes in behavior was much larger.

The results have important implications for the interpretation of ecological studies about the effects of policy interventions, including school closures. The finding of widespread behavioral change before school closures took effect suggests that both school closures and behavioral change have a common cause: a general fear of contracting and spreading disease.

These results add clarity to the existing literature by testing the independent role of legal restrictions of school closures and bans on large gatherings when controlling for contemporaneous voluntary behavioral change. Unlike the results of Courtemanche et al., this analysis finds a significant and beneficial association of school closures with pandemic spread,
Voluntary behavior is connected to legal restrictions in complex ways. A previous study found evidence of asymmetric results between closures and reopening, with reopening appearing to be considered an all-clear: although there were modest independent effects of stay-at-home orders and closures of restaurants, reopening restaurants was associated with a surge in cases. Therefore, as policy makers plan for school reopening, it is imperative that they do so in a way that does not signal to residents that they can safely return to pandemic behaviors. Proper communication that school reopenings are potentially hazardous but worthwhile for child well-being is one such way to achieve this aim.

### Table. Association Between Logged 7-Day Rolling Mean Daily Outcomes and School Closure, Gathering Bans, and Behavioral Change at State Level*

<table>
<thead>
<tr>
<th>Measure</th>
<th>OLS models</th>
<th>Prais-Winsten models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>School closure only</td>
<td>School closure and gathering ban</td>
</tr>
<tr>
<td></td>
<td>(model 1)</td>
<td>(model 2)</td>
</tr>
<tr>
<td>Daily incidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days into study period</td>
<td>1.151 (1.145-1.158)</td>
<td>1.147 (1.140-1.154)</td>
</tr>
<tr>
<td>Days since 14-d lag, IRR (95% CI)</td>
<td>0.877 (0.870-0.883)</td>
<td>0.916 (0.907-0.924)</td>
</tr>
<tr>
<td>School closure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gathering ban</td>
<td>0.956 (0.949-0.963)</td>
<td>NA</td>
</tr>
<tr>
<td>15% Reduction in time spent at work</td>
<td>NA</td>
<td>0.860 (0.853-0.866)</td>
</tr>
<tr>
<td>No. of observations</td>
<td>3000</td>
<td>3000</td>
</tr>
<tr>
<td>R² value</td>
<td>0.643</td>
<td>0.662</td>
</tr>
<tr>
<td>Daily death rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days into study period</td>
<td>1.165 (1.158-1.172)</td>
<td>1.162 (1.155-1.169)</td>
</tr>
<tr>
<td>Days since 21-d lag, IRR (95% CI)</td>
<td>0.858 (0.850-0.866)</td>
<td>0.899 (0.887-0.909)</td>
</tr>
<tr>
<td>School closure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gathering ban</td>
<td>0.956 (0.942-0.962)</td>
<td>NA</td>
</tr>
<tr>
<td>15% Reduction in time spent at work</td>
<td>NA</td>
<td>0.846 (0.839-0.854)</td>
</tr>
<tr>
<td>No. of observations</td>
<td>3000</td>
<td>3000</td>
</tr>
<tr>
<td>R² value</td>
<td>0.686</td>
<td>0.695</td>
</tr>
</tbody>
</table>

Abbreviations: IRR, incidence rate ratio; NA, not applicable; OLS, ordinary least squares.

* Estimates are retransformed from the log scale by exponentiation and should therefore be interpreted as IRRs. In addition to the covariates shown, all models additionally control for natural log population, percentage of the population younger than 15 years, percentage of the population older than 65 years, percentage of the population in nursing facilities, quartile of cumulative incidence growth in the first 10 days of the study period, state-level social vulnerability index, daily testing per 1000 population, rate of obesity, and urban density. The full model results are available in eTables 9 and 10 in the Supplement.

**Days since** represents how much earlier in the pandemic either school closures, gathering bans, or work reductions happened. Each 1-unit increase in a “days since” variable represents 1 day earlier that the change happened.

### Limitations
This study has some limitations. Data on behavior are from indirect observation using cell phone locations. Although this method has been previously used in the literature and has good validity, it is subject to measurement error, and it is not reported at geographic levels more granular than the state for privacy reasons. The publicly available data therefore miss potentially important within-state variation in behavioral change. We were unable to find any formal study of the validity of Google’s identification of work or home locations.

Local school closures happened in some areas earlier than statewide closures of schools. This analysis, like previous analyses, uses the statewide closures. If a substantial number of schools closed earlier in many states, it may undermine the conclusion that behavioral change happened earlier than school closures.

It is essential to be mindful of the limitations of the simulations. Because the growth in COVID-19 cases is exponential, modest changes in the early stages of the epidemic are modeled as having enormous effects over time. Therefore, extrapolating beyond the range of data is risky. Modeled results herein are presented to provide context to the empirical results, but

albeit of modest magnitude. Unlike the results of Auger et al., this analysis finds a significant and beneficial association of bans on large gatherings with pandemic spread, albeit of modest magnitude. These differences may result from better control for serial autocorrelation and collider bias. It seems plausible that the timing of any given legal restriction is influenced by both prior rates of severe acute respiratory syndrome coronavirus 2 transmission as well as by other legal restrictions in place and, if so, each legal restriction is a collider of others.

Voluntary behavior is connected to legal restrictions in complex ways. A previous study found evidence of asymmetric results between closures and reopening, with reopening appearing to be considered an all-clear: although there were modest independent effects of stay-at-home orders and closures of restaurants, reopening restaurants was associated with a surge in cases. Therefore, as policy makers plan for school reopening, it is imperative that they do so in a way that does not signal to residents that they can safely return to pandemic behaviors. Proper communication that school reopenings are potentially hazardous but worthwhile for child well-being is one such way to achieve this aim.
these—and all similar exercises—should be interpreted with caution. If school closures had in fact happened 2 weeks later, the resulting surge in cases may have triggered stronger voluntary behavioral changes that would have blunted the effect of keeping schools open, leading to fewer than the estimated additional 23,000 deaths.

Conclusions
School closures are a blunt instrument for achieving reductions in the spread of any pandemic. They cause widespread disruption to the economy, including to the health care workforce, and they are associated with substantial harms to child development. Individuals in the US were changing their behavior in large numbers before they were compelled to do so by official policy, suggesting it is possible to achieve the benefits of a social distancing strategy without incurring costs to the most vulnerable members of society.

Given the uncertainty surrounding the severity and transmission dynamics of the disease, the decision to close schools in spring 2020 was reasonable. However, this analysis suggests that school closures were not the factor most strongly associated with slowing the spread of the disease. Other, less harmful, means may be found in the future to effectively limit further spread of COVID-19.
The timing of school closures and behavioral changes with COVID-19 in the US

REFERENCES


