medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

College Openings, Mobility, and the Incidence of COVID-19

Martin S. Andersen[a,1,2], **Ana I. Bento**[b,1,2], **Anirban Basu**[c], **Christopher R. Marsicano**[d,e], **and Kosali Simon**[f]

[a]Department of Economics, University of North Carolina at Greensboro, Greensboro, NC 27412; [b]Department of Epidemiology and Biostatistics, School of Public Health, Indiana University-Bloomington, Bloomington, IN 47405; [c]The Comparative Health Outcomes, Policy, and Economics (CHOICE) Institute, Departments of Pharmacy, Health Services, and Economics, University of Washington, Seattle, Seattle, WA 98195; [d]The College Crisis Initiative at Davidson College, Davidson College, Davidson, NC 28035; [e]Educational Studies Department, Davidson College, Davidson, NC 28035; [f]O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington, IN 47405

School and college reopening-closure policies are considered one of the most promising non-pharmaceutical interventions for mitigating infectious diseases. Nonetheless, the effectiveness of these policies is still debated, largely due to the lack of empirical evidence on behavior during implementation. We examined U.S. college reopenings' association with changes in human mobility within campuses and in COVID-19 incidence in the counties of the campuses over a ten-week period around college reopenings. We used an integrative framework, with a difference-in-differences design comparing areas with a college campus, before and after reopening, to areas without a campus and a Bayesian approach to estimate the daily reproductive number (R_t). We found that college reopenings were associated with increased campus mobility, and increased COVID-19 incidence by 2.7 cases per 100,000 (95% confidence interval [CI]: 1.4 – 4.0), of which 0.042 per 100,000 resulted in death (95% CI: 0.002 – 0.082). This reflected our estimate of increased transmission locally after reopening. A greater increase in county COVID-19 incidence resulted from campuses that opened for in-person (vs. online) teaching and with greater exposure to students from counties with high COVID-19-incidence in the weeks before reopening. Our study sheds light on movement and social mixing patterns during the closure-reopening of colleges and offers strategic instruments for benefit-cost analyses of future school reopening/closure policies.

COVID-19 | college reopening | disease transmission| community spillover | mobility and mixing

One of the key lessons learned from the COVID-19 pandemic has been the pivotal role of mobility and mixing in spreading the infection. These phenomena are acutely important in congregate and communal living settings that are common in nursing homes, prisons, and colleges in the United States and globally (1–4). However, the role of communal living, and its interaction with mobility and mixing, is difficult to identify empirically since people enter communal living settings non-randomly. The resumption of teaching on a college campus provides a sudden change in a community's exposure to communal living and differences across college campuses lead to variation in the extent to which campus reopenings induce mixing with higher and lower incidence areas.

The susceptibility of children and college-age individuals to COVID-19 and their role in transmission has been heavily debated and remains hard to quantify (5–9). Following the first wave of school closures in the United States, COVID-19 incidence fell across the country, leading many public health officials to view closing schools as a viable strategy to mitigate the spread of the pandemic (e.g., 10, 11). However, closing schools, while potentially reducing the spread of COVID-19, may adversely affect children and college students. As a result, it is important to understand what role college reopenings play, if any, in the COVID-19 pandemic to design efficient mitigation strategies now and in the future.

During late Summer 2020, colleges and universities across the United States reopened and welcomed hundreds of thousands of students back to campus in the United States (12). Over half of these institutions reopened for in-person teaching, although many institutions switched to online instruction after rapid increases in reported COVID-19 cases on campuses and in the community (13, 14). A handful of studies have sought to formally test the hypothesis that reopening college campuses increased COVID-19 incidence (4, 15–18). However, the institutions in these studies represent a small proportion of the 11 million undergraduates enrolled in public and non-profit four-year institutions across the country (12). A phylogenetic study from western Wisconsin (3) identified two clusters of SARS-CoV-2 strains on college campuses that may have subsequently infected nursing home residents, demonstrating transmission between college campuses and the surrounding community. Nonetheless, the effectiveness of college reopening policies as non-pharmaceutical interventions for mitigating the burden of COVID-19 is still disputed. Simulation-based studies have been unable to provide public health officials with conclusive recommendations, despite detailed COVID-19 transmission datasets (e.g., 19, 20)). The lack of a clear direction is mostly

DRAFT

Significance Statement

College campuses represent a challenge for COVID-19 mitigation efforts. Mixing behavior associated with communal living, nationwide student migration, and few assessments of spillover effects on surrounding communities make reopening campuses for in-person instruction controversial. Our study uses mobility as a proxy for migration and mixing behavior in campus communities before and after reopening. Using college reopening plans and cell phone mobility data, we evaluate county-level age-structured COVID-19 transmission evolution and mortality due to campus teaching modality. We find substantial increases in mobility and cases across all modalities and increased deaths associated with primarily in-person instruction. These findings suggest leaders of communal living environments should plan around COVID-19 incidence in their vicinity and the areas from which they draw community members.

Author contributions: M.A., A.I.B., A.B., C.M., and K.S. designed the research and wrote the paper. M.A. and A.I.B. performed the research and analyzed the data. C.M. contributed novel data.

The authors declare no competing interests.

[1]M.A.(Author One) contributed equally to this work with A.I.B (Author Two).

[2]To whom correspondence should be addressed. E-mail: msander4@uncg.edu or abento@iu.edu

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

due to a lack of data about the college-specific details and how to harness movement as proxies for behavior and mixing patterns of the population while such strategies are in place.

We harnessed comprehensive, national data covering the start date and instructional method of most four-year U.S. colleges and universities together with a highly resolved dataset (both spatially and by age) from the CDC, (21) which provided detailed demographic information on COVID-19 cases around the country. This gave us the ability to directly measure the variation in human movement patterns caused by the policy and, in addition, allowed us to identify college-age cases and assign cases based on symptom onset. We hypothesized that reopening colleges would increase COVID-19 transmission within the college community with potential spillover effects onto the neighboring populations. We also hypothesized that increases in incidence would be greater on campuses that attract students from areas with a higher incidence of COVID-19 and that these effects would be concentrated among campuses providing face-to-face instruction. While there is some compelling research around testing (22, 23) and limiting student mobility (9) as COVID-19 mitigation strategies, it is outside of the scope of our study to understand the impact of specific actions colleges may have taken in response to rising rates.

We use an integrative framework, with a difference-in-differences design comparing areas with a college campus, before and after reopening, to areas without a campus and a Bayesian approach to estimate the daily reproductive number (R_t). We unequivocally demonstrate that there was a marked increase in COVID-19 incidence among college-age students following the reopening of campuses. Finally, while COVID-19 case counts have been a focus of several studies, our data also allowed us to examine other public health outcomes, such as hospitalizations or deaths. Our results provide evidence of the COVID-19 impact of colleges-reopening policies locally and in neighboring communities.

Results

Our study period ran from June 24th 2020 to November 9th 2020, which spanned the five weeks before the first campus reopening and five weeks after the last campus reopened. Of the 3,142 counties of the United States, 784 contained a college campus from our universe of 1,360 colleges. However, over 238.0 million people live in counties with a college campus. Figure S1 of the Supplementary Information maps the campuses in our sample by teaching modality.

Our identification strategy made comparisons between counties with and without a college campus around the time that a campus reopened in a "difference-in-differences" design (24). Since several counties contained more than one college campus, we weighted each campus by its share of total college enrollment in the county and explored the sensitivity of our results to the assignment of reopening dates in the Supplementary Information. Based on preliminary results on mobility, which indicated that on-campus mobility increased significantly the week before the resumption of classes, we defined our "post" period as beginning one week before classes resumed.

Event studies. The reopening of a specific college affected mixing patterns not only of the students but also the members of the surrounding communities where these students live. The number of devices on campus increased significantly in the week before campuses reopened and remained high for at least the first 14 days following reopening (Figure 1a). Aggregating by week, which smooths out day of the week fluctuations in movement, and separating the sample by teaching modality demonstrated that there were significant increases in movement to census block groups containing college campuses after those campuses reopened regardless of the teaching modality (Figure 1b), although the increase was larger for in-person reopenings. The increase in mobility was accompanied by a rise in COVID-19 incidence, with new cases assigned to the date of symptom onset (Figure 1c). The increase in COVID-19 cases was accompanied by an increase in cases requiring hospitalization and that resulted in death, particularly among campuses that opened for primarily online teaching(Figure 1, panels (d) and (f)). R_t increased regardless of the teaching modality chosen by the college (Figure 1g).

Difference-in-differences estimates. The reopenings lead to a cascade of indirect effects at the population level. To show this, we estimated a series of difference-in-difference models to estimate the effect of reopening a college campus on mobility and COVID-19 outcomes. We present the detailed results in the Supplementary Information (Table S1), but describe the results below.

Reopening a college campus was associated with a 41.3 log point increase in the number of devices on campus, or approximately a 51.1% increase in movement on campus, from the week prior to the start of classes (Table S1, column (1)). The increase in movement was larger in schools that reopened for primarily in-person, as opposed to primarily online, instruction (47.1 vs. 33.0 log points or 60% vs. 39%, $p < 0.001$) and began increasing the week before reopening. The increase in mobility was larger for colleges that had greater exposure to students from areas with high levels of COVID-19 incidence.

Using our difference-in-difference framework, we found that reopening a college was associated with a statistically significant increase of 2.7 cases per 100,000 people (Table S1, column (2)). The increase in COVID-19 incidence was larger in areas that were exposed to areas with high COVID-19 incidence (an additional 1.6 cases per 100,000 people per 10000 devices arriving in the county), while counties with campuses that reopened for primarily online education experienced a smaller increase in daily COVID-19 incidence, compared to those that opened for in-person instruction (1.7 vs. 3.2 cases per 100,000, respectively, $p = 0.087$). Consistent with our mobility data, COVID-19 incidence became progressively larger from the week before classes resumed and peaked during the first two weeks after classes (1.4 and 3.7 cases per 100,000, respectively) resumed before declining to a lower, but still elevated, level of 2.5 cases per 100,000 per day.

Reopening a college campus was also associated with a marginally significant ($p < 0.1$ level) increase of 0.053 cases per 100,000 that ultimately require hospitalization (column 3) and 0.013 cases per 100,000 that needed ICU admission (column 4). The increase in cases needing hospitalization or ICU admissions was associated ($p < 0.1$) with reopening a campus for in-person, but not online instruction.

We also identified an increase in cases resulting in deaths (column 5) by the end of December, with an additional 0.04 COVID-19 deaths per 100,000 people following reopening and a larger 0.05 per 100,000 increase in COVID-19 cases resulting in death for campuses that reopened for in-person education.

DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

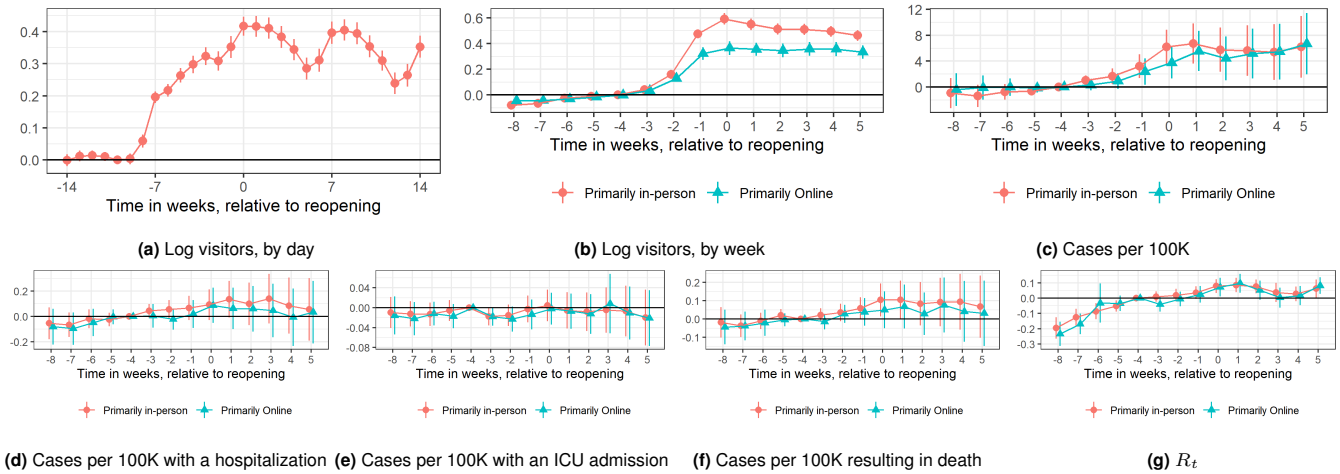


Fig. 1. Event study estimates of reopening college campuses, relative to counties without a college campus. Colors refer to teaching modality, where red is primarily in person and blue is primarily online. (a) Reopening college campuses significantly increased the number of devices on college campuses, with evidence of a steady increase in the ten days before the start of classes. These increases persisted on college campuses for at least six weeks after reopening (b) and was larger for campuses that reopened for primarily in-person teaching. Reopening college campuses increased new cases per 100,000 in the subsequent weeks for both teaching modalities (c), with a significant difference between teaching modality for the week classes resumed. Estimates for the number of new cases resulting in hospitalizations (d), ICU admissions (e), or mortality (f) were noisy, but indicated that in-person reopenings appear to be associated with an increase in hospitalizations and deaths. Local transmission, measured by R_t, increased after reopening a college, regardless of teaching modality (g). Across all models, except for R_t, there was little evidence of a trend leading up to roughly three weeks before campuses reopened, which is when mobility began to increase on college campuses.

The central epidemiological parameter governing a disease system's dynamics is the effective reproduction number (R_t). We estimated a significant increase in daily (R_t) around the time of reopening, consistent with an uptick in transmission. On average there was an increase in R_t of 0.164 (CI: 0.137 - 0.191). We note that R_t did not significantly differ by the teaching method chosen for the campus (Table S1).

Age-specific incidence. To observe the age-stratified dynamics, we explored age-specific incidence. Our analyses supported the conclusion that the shifts in age dynamics overtime likely resulted from college reopenings in Figure 2. The top panel (Figure 2a) demonstrates a clear shift, where we observe an increase in COVID-19 incidence in people ages 10 - 29, while incidence did not increase appreciably in any other age group. However, the increase in these college-aged students was dramatic, with incidence increasing by almost seven additional cases per 100,000 people between 10 and 29 years of age. The second panel indicates that our estimates of the effect of reopening on hospitalizations by age group are too noisy to draw any inferences. In the third panel, we demonstrate that reopening a college campus increases the number of incidence COVID-19 cases that result in an ICU admission by the end of 2020. The fourth panel demonstrates that the overall increase in mortality we observed was not isolated to any particular age group, although there was a small reduction in incident cases among 0 to 9 year-olds resulting in death.

While the data appear to paint a clear picture, it is possible that several mechanisms may yield a similar age-specific profile of cases. Thus, in the SI, we test these observations. We show age-specific event studies (Figure S2) for our four age-specific outcomes. These event studies clearly demonstrate that COVID-19 incidence rose beginning in the weeks immediately before campuses reopened and remained elevated subsequently, with statistically significant increases after 4 or more weeks for children under 10 years of age as well as people 70 and older. The event studies for hospitalization, ICU admissions, and death typically present little evidence of a time trend in advance of the campus reopening, although we also find few other statistically significant changes in our outcomes.

Differential effects by teaching modality. We separated teaching modality into more detailed categories to further explore our finding that in-person teaching was more strongly associated with increases in COVID-19 cases than online teaching. Our results demonstrate that campuses that reopened with a greater emphasis on in-person teaching were associated with larger increases in mobility and incident COVID-19 cases (Figure 3), with a slight increase in the number of cases per 100,000 that resulted in ICU admission and death.

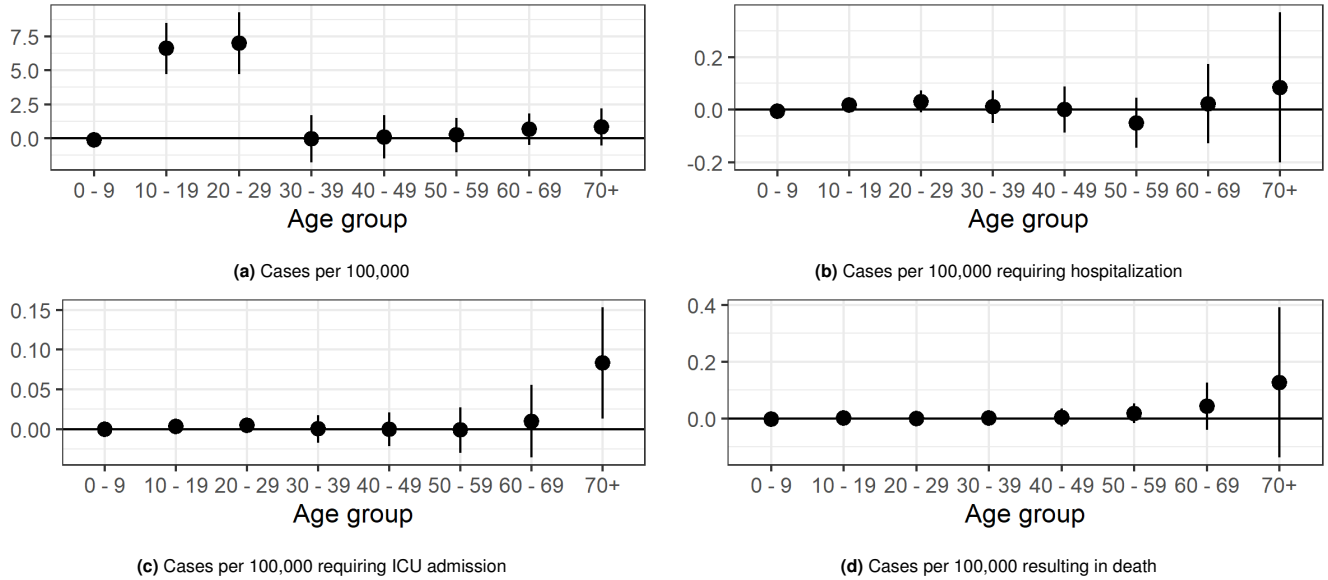
Robustness and alternative specifications. We considered a number of alternative models, which we report fully in the Supplementary Information (section 2). Our results were robust to using a trimmed, as opposed to a balanced, panel, restricting to counties with a college, using a single college in each county, and introducing quadratic time trends. However, our results were sensitive to weighting counties by population, or omitting county-specific time trends or date fixed effects.

Discussion

Our results provide a quantitative evaluation of mobility patterns during periods of reactive college closure and reopening strategies and highlight their impact in shaping social interactions of the college and surrounding communities. We found that college policies induced marked changes in the overall number of daily mobility interactions. Our findings demonstrate that re-opening a college was associated with an increase in the number of cellular devices on campus (a dramatic increase in population size) leading up to the start of classes and after classes resumed for all teaching modalities, although

DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



(a) Cases per 100,000

(b) Cases per 100,000 requiring hospitalization

(c) Cases per 100,000 requiring ICU admission

(d) Cases per 100,000 resulting in death

Fig. 2. Age-specific effects of college reopenings. (a) demonstrates that the increase in the incidence of COVID-19 was isolated to people between 10 and 29 years of age, which encompasses college-age individuals. There were, however, no statistically significant and age-specific changes in the incidence of cases resulting in hospitalization (b). Nevertheless, the incidence of cases resulting in an ICU admission rose by almost one additional case per million people 70 and older (c). We did not find any age-specific increases in mortality due to COVID-19, although these results were imprecisely estimated (d). Point estimates and standard errors are available in Table S2 of the Supplementary Information.

DRAFT

the increase in mobility is larger for in-person as opposed to online teaching. We unequivocally showed that re-opening a college significantly increased the incidence of COVID-19 in the county as well as increased COVID-19 mortality, with marginally significant increases in hospitalizations and ICU admissions. In general, these increases were larger or only present for counties that contain more universities that re-opened for in-person, as opposed to online, education, and persisted after reopening. We also demonstrated that counties containing colleges that drew students from areas with higher COVID-19 incidence experienced significantly larger increases in COVID-19 incidence following campus reopening. This is likely induced by the dramatic increase in the number of contacts of students with each other on the campuses and with the surrounding communities.

To contextualize our findings, there are 238.0 million Americans in the 784 counties that contain a college campus in our sample. Our results demonstrate that reopening college campuses resulted in an additional 6,500 (6492 [95% CI: 3381 – 9604]) cases of COVID-19 per day. This estimate is consistent with the aggregate number of cases reported on the New York Times case tracker which reported more than 397,000 cases as of December 11 2020 (25), which would correspond to nine weeks of additional cases using our main estimate.

However, because of the nature of the cases reports data, we were unable to disentangle how many of the cases we measure as our outcome are "imported" (student arrivals) and how many are local transmissions from the students. Further, asymptomatic cases were only identified if testing was done on campus regardless of symptoms. Nevertheless, our results are inconsistent with large numbers of "imported" cases since an imported case would lead to an increase in COVID-19 cases contemporaneously with any increase in mobility, while we observed a one-week lag between peek mobility and the peek change in COVID-19 incidence, when cases are assigned based on symptom onset.

We did not quantify potential spillovers to the communities surrounding campuses, as these effects would require college-level incidence data, which are not consistently collected. However, using age-specific data, we were able to demonstrate that most of the increase in COVID-19 incidence arose among college-aged students (ages 10-29), with an increase in incidence requiring ICU admission among people 70 and older as well.

Additional work is necessary to identify the optimal reopening- closure policies (e.g., lengths) and under which circumstances specific policies are cost-effective. However, evaluating the effectiveness of specific mitigation measures taken by colleges, especially the ways in which colleges have reacted to the initial increases in cases with strong countermeasures, was beyond the scope of this initial study and remain priorities for future studies. Similarly, we were unable to test what has occurred once colleges change decisions, such as changing instructional modes temporarily or encouraging students to return home (9) since these changes were reactions to rapidly increasing case counts (26).

While we only directly demonstrate that college campuses that were more heavily exposed to COVID-19 lead to larger increases in incidence, our results also indicate that sending students home from colleges due to high COVID-19 incidence is likely to lead to increased COVID-19 incidence in students' home communities since the same exposure mechanism would run in reverse. Public health officials have also raised these concerns, some of whom have publicly opposed closing dormitories, even after a college or university transitioned to online education (27). Further research on the effects of sending students home is needed to understand the risks and benefits of closing residence halls.

The nature of our data limits our results. Our mobility analysis relies on observing cellular GPS signals and these

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

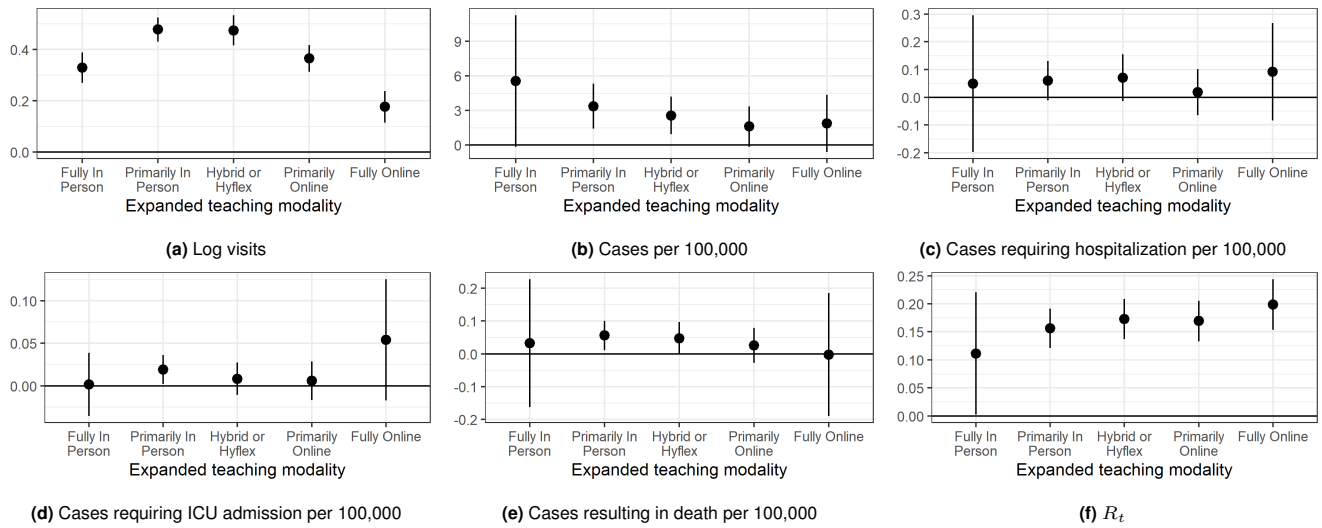


Fig. 3. Differential effects of reopening college campuses by expanded teaching modality. (a) Campuses that reopened for "Primarily in-person" and "Hybrid" teaching had the largest increase in devices on campus following reopening, while the increase in visitors was significantly smaller for fully online reopenings. All reopenigs except "Fully Online" were associated with a significant increase in COVID-19 cases after reopening (b). There were no statistically significant effects of reopening a college campus by teaching modality for hospitalizations (c) or ICU admissions (d). However, there was a statistically significant increase in mortality following a reopening that was "primarily in person" (e). Except at the extremes, for which we had relatively few colleges, there were uniform increases in R_t, regardless of the teaching modality (f).

devices may not always report their location. In addition, it is unlikely that devices correspond in a one-to-one manner with people since college students may have more than one device (a phone and a cell-enabled tablet) that provide data under distinct identifiers. Second, we are unable to measure cases among college students vs. others in the county community, beyond using the age of the individual. Third, our mobility measure does not take account of students who may live in off-campus housing and take classes online.

Our results demonstrate the essential role that mixing and mobility play in seeding COVID-19 in the community and the role that congregate living settings play in providing a fertile ground for COVID-19 to expand. For example, these results highlight the role that nursing homes and prisons play in the COVID-19 pandemic and complement existing research on cross-nursing home linkages and COVID-19 incidence (1).

Our analysis is a good step towards building a framework to map mobility to contacts (as COVID-19 era contact matrices become available (28)), in the analysis of airborne infectious diseases. As such, our framework has a much wider scope than the study of COVID-19 related college policies in one specific region. Our findings are critical in the context of adapting public health management strategies, as they consider additional strategies to mitigate disease burden and decrease transmission. The effects of college reopenings are also informative for outbreak management in other communal settings, including nursing homes and prisons, both of which have been particularly hard hit by the COVID-19 pandemic.

Despite all data limitations, our study, provides (i) empirical evidence about changes in "behavior" (mobility surges) of the population during the implementation of the school-reopening strategies, (ii) a multi pronged approach to estimate mobility patterns and evaluate effects on the spread of infectious diseases with an unique degree of detail and (iii) tools for evidence-based decision-making beyond evaluating college reopening strategies.

Materials and Methods

Study Data. We collected data on opening dates and announced instructional methods from the College Crisis Initiative at Davidson College (C2i) (29) for 1,431 public and non-profit colleges and universities ("colleges") in the United States. The College Crisis Initiative collects data on nearly all non-profit and public four-year degree-granting institutions with full-time undergraduates that receive Title IV aid. It excludes four-year for-profit institutions, specialty institutions like seminaries or stand-alone law schools, or institutions with graduate-only programs. This list comes from the Integrated Postsecondary Education Data System (IPEDS), which lists in total 6,527 institutions ranging from research universities to non-degree-granting institutions like local cosmetology schools. IPEDS indicates that of those, 2,009 are four-year public and non-profit degree-granting institutions with first-time, full-time undergraduates. Our sample, therefore, represents nearly 70 percent of these institutions. Further, this represents 70 percent of total undergraduate enrollment among all institutions of higher education in the United States (author calculations based on IPEDS administrative 2018 data).

We assigned college campuses to Census Block Groups (CBGs) using a college campus shapefile (geographic coordinates) prepared by the Department of Homeland Security (30). We used a spatial join to assign each Census Block Group to the college campus that occupied the largest area in the block group, as a result, our assignment of campuses to block groups was unique. We then merged these data with college opening dates. Our final sample included 1360 schools in 1398 college-county pairs. We assigned reopening strategies based on the mode of instruction reported on the date instruction began for Fall 2020. Campuses were classified as primarily in-person or primarily online based on the instructional modality in effect the day classes resumed for the Fall semester. Institutions that instituted primarily hybrid (376) or in-person (493) modes of instruction were classified as "primarily in person." Institutions that offered only online classes, or for which the majority of the classes offered were online were classified as "primarily online" (490).

We extracted cellular data from SafeGraph's Social Distancing Metrics files. SafeGraph aggregates anonymized location data from numerous applications in order to provide insights about physical places, via the Placekey Community. To enhance privacy, SafeGraph excludes CBG information if fewer than five devices visited an establishment in a month from a given CBG. These data measure

DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

the number of devices that are detected each day in each CBG, from June 24th through November 11th. SafeGraph data have been used in several recent publications (31–36).

We used deidentified, case-level data from the Centers for Disease Control to estimate the incidence of COVID-19 in a county by age-group (21).* These data are extracted from the CDC "Human Infection with 2019 Novel Coronavirus Case Report Form", which we have replicated in the Supplementary Information. The CDC data include several possible dates for each case we used, in order of priority, the date a sample was collected, the date of symptom onset, and the date a case was reported to the CDC. We then counted the number of cases diagnosed in each county, age-group, date cell and the number of cases that were, by December 31, 2020, hospitalized, admitted to the ICU, or resulted in death. We converted these values into values per 100,000 people in a age-county cell using population data from the 5-year American Community Survey (37). The CDC data identifies the ultimate outcome of cases by diagnosis date, so our data indicates the number of incident cases per 100,000 people and the number that resulted in death, hospitalization, or ICU admission.

In the Supplementary Information, we describe our estimation of the effective reproductive number ($R(t)$) and how we constructed our index for college exposure to other counties.

All data used can be requested from SafeGraph, the CDC, and the College Crisis Initiative.

Statistical Analysis. Our main analyses use a balanced panel of counties and Census Block Groups (CBGs). In cases where a county has multiple colleges represented (as happens for 266 of 3139 counties), we weighted each college in the county by its share of the county's total college enrollment.

We used event-study and difference-in-difference methods to assess the relationship between college reopenings and our main outcomes: mobility to campuses, COVID-19 incidence, and COVID-19 cases resulting in a hospitalization, ICU admission, or death by December 31 2020.

Our event study assessed the changes in mobility and changes in COVID-19 cases, hospitalizations, ICU admissions, and deaths, relative to when a college has reopened, controlling for the geographic level of our analysis—either the CBG (for the mobility models) or county (for the COVID-19 models)—calendar date, and college (since a college can span more than one CBG or county).

Our regression models for the event studies, using i to denote the geographic unit (CBG or county), c the college in that unit, and t to denote time can be written as:

$$O_{it} = \sum_{\tau=-14,\tau\neq-10}^{14} \beta_\tau 1[t - R_c = \tau] + \delta_i + \sigma_t + \chi_c + \eta_i \times t + \varepsilon_{ict}$$

Where O_{it} the outcome, R_c is the reopening date for campus c, δ_i, σ_t, and χ_c are county, time, and college fixed effects, $\eta_i \times t$ is a set of county-specific time trends, and ε_{ict} is an idiosyncratic error term. We omit the county-specific time trends when using R_t as the dependent variable since R_t. We also estimated our event study model using week indicators for the period beginning 8 weeks prior to reopening and ending five weeks after reopening (the fourth week before reopening is the omitted reference level) to smooth out daily fluctuations due to day-of-week effects and provide a longer horizon to visually detect pre-trends. For our difference-in-difference models we replaced the time relative to opening indicators with a single indicator for the post period (beginning one week before reopening to accommodate students returning to campus).

We assessed changes relative to the reopening date in all models, controlling for county (or CBG), college, and date effects. We also estimated models that included interactions with an indicator for a campus being primarily in-person, our student exposure index, and breaking up the post period into the week before the start of classes, weeks 0 and 1, weeks 2 and 3, and weeks 4+. To incorporate variation across age groups, we also estimated age-specific event studies and difference-in-difference models.

*The CDC does not take responsibility for the scientific validity or accuracy of methodology, results, statistical analyses, or conclusions presented.

References

1. Chen MK, Chevalier JA, Long EF (2021) Nursing home staff networks and covid-19. *Proceedings of the National Academy of Sciences* 118(1).
2. Kinner SA, et al. (2020) Prisons and custodial settings are part of a comprehensive response to covid-19. *The Lancet Public Health* 5(4):e188–e189.
3. Richmond CS, Sabin AP, Jobe DA, Lovrich SD, Kenny PA (2020) Sars-cov-2 sequencing reveals rapid transmission from college student clusters resulting in morbidity and deaths in vulnerable populations. *medRxiv*.
4. Leidner AJ (2020) Opening of large institutions of higher education and county-level covid-19 incidence—united states, july 6–september 17, 2020. *MMWR. Morbidity and Mortality Weekly Report* 70.
5. Auger KA, et al. (2020) Association between statewide school closure and covid-19 incidence and mortality in the us. *Jama* 324(9):859–870.
6. Laws RL, et al. (2021) Symptoms and transmission of sars-cov-2 among children—utah and wisconsin, march–may 2020. *Pediatrics* 147(1).
7. Kim J, et al. (2020) Role of children in household transmission of covid-19. *Archives of disease in childhood*.
8. Ludvigsson JF (2020) Children are unlikely to be the main drivers of the covid-19 pandemic–a systematic review. *Acta Paediatrica* 109(8):1525–1530.
9. Mangrum D, Niekamp P (2020) Jue insight: College student travel contributed to local covid-19 spread. *Journal of Urban Economics* p. 103311.
10. Markel H, et al. (2007) Nonpharmaceutical interventions implemented by us cities during the 1918-1919 influenza pandemic. *JAMA* 298(6).
11. Yehya N, Venkataramani A, Harhay MO (2020) Statewide Interventions and Covid-19 Mortality in the United States: An Observational Study. *Clinical Infectious Diseases*. ciaa923.
12. Marsicano C, et al. (2020) C2i fall 2020 dataset. [data file and code book].
13. Nierenberg A, Pasick A (2020) Schools briefing: University outbreaks and parental angst. *The New York Times*.
14. Hubler S, Hartocollis A (2020) How Colleges Became the New Covid Hot Spots. *The New York Times*.
15. Stubbs CW, Springer M, Thomas TS (2020) The impacts of testing cadence, mode of instruction, and student density on fall 2020 covid-19 rates on campus. *medRxiv*.
16. Lu H, et al. (2021) Are college campuses superspreaders? a data-driven modeling study. *Computer Methods in Biomechanics and Biomedical Engineering* 0(0):1–11. PMID: 33439055.
17. Li Y, et al. (2021) Association of university reopening policies with new confirmed covid-19 cases in the united states. *medRxiv*.
18. Badruddoza S, Amin MD (2020) Causal impacts of teaching modality on us covid-19 spread in fall 2020 semester. *medRxiv*.
19. Aleta A, Martín-Corral, D., Pastore y Piontti A, Moreno Y (2020) Modelling the impact of testing, contact tracing and household quarantine on second waves of covid-19. *Nature Hum Behav* 4:964–971.
20. Liu QH, et al. (2021) The covid-19 outbreak in sichuan, china: Epidemiology and impact of interventions. *PLOS Computational Biology* 16(12):1–14.
21. Centers for Disease Control and Prevention, COVID-19 Response (2020) Covid-19 case surveillance data access, summary, and limitations (version date: December 31, 2020).
22. Denny TN, et al. (2020) Implementation of a pooled surveillance testing program for asymptomatic sars-cov-2 infections on a college campus—duke university, durham, north carolina, august 2–october 11, 2020. *Morbidity and Mortality Weekly Report* 69(46):1743.
23. Paltiel AD, Zheng A, Walensky RP (2020) Assessment of sars-cov-2 screening strategies to permit the safe reopening of college campuses in the united states. *JAMA network open* 3(7):e2016818–e2016818.
24. Goodman-Bacon A (2018) Difference-in-Differences with Variation in Treatment Timing, (National Bureau of Economic Research), Working Paper 25018.
25. The New York Times (2021) *Coronavirus (Covid-19) Data in the United States*.
26. Moon S (2020) This california university is telling students to vacate dorms just 1 week after starting classes (https://www.cnn.com/world/live-news/coronavirus-pandemic-08-31-20-intl/index.html).
27. Yan H, Fox M, Gumbrecht J (2020) CDC official affirms coronavirus deaths really are coronavirus deaths (https://www.cnn.com/2020/09/02/health/us-coronavirus-wednesday/index.html).
28. Prem K, et al. (2020) Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the covid-19 era. *medRxiv*.
29. Marsicano C (2020) C2I: The College Crisis Initiative – Crisis to Innovation (https://collegecrisis.org/).
30. Department of Homeland Security (2020) Colleges and Universities Campuses.
31. Weill JA, Stigler M, Deschenes O, Springborn MR (2020) Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proceedings of the National Academy of Sciences* 117(33):19658–19660.
32. Holtz D, et al. (2020) Interdependence and the cost of uncoordinated responses to COVID-19. *Proceedings of the National Academy of Sciences*.
33. Andersen M (2020) Early Evidence on Social Distancing in Response to COVID-19 in the United States, (Social Science Research Network, Rochester, NY), SSRN Scholarly Paper ID 3569368.
34. Andersen M, Maclean JC, Pesko MF, Simon KI (2020) Effect of a Federal Paid Sick Leave Mandate on Working and Staying at Home: Evidence from Cellular Device Data, (National Bureau of Economic Research), Working Paper 27138.
35. Gupta S, et al. (2020) Tracking Public and Private Response to the COVID-19 Epidemic: Evidence from State and Local Government Actions, (National Bureau of Economic Research), Working Paper 27027.
36. Nguyen TD, et al. (2020) Impacts of State Reopening Policy on Human Mobility, (National Bureau of Economic Research), Working Paper 27235.
37. Ruggles S, et al. (2020) IPUMS USA: Version 10.0.

DRAFT

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

PNAS
www.pnas.org

Supplementary Information for

College Openings, Mobility, and the Incidence of COVID-19

Martin S. Andersen, Ana I. Bento, Anirban Basu, Christopher R. Marsicano, Kosali Simon

Martin S. Andersen.
E-mail: msander4@uncg.edu

This PDF file includes:

Supplementary text
Figs. S1 to S4 (not allowed for Brief Reports)
Tables S1 to S3 (not allowed for Brief Reports)
SI References

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Supporting Information Text

1. Additional Methods

Effective reproductive number ($R(t)$). We used the daily cases, incubation period, and serial interval previously estimated (1). This allowed us to estimate the effective reproduction number for each county. The effective reproduction number $R(t)$ represents the mean number of secondary cases generated by a primary infector at time t (2–5). This measure is useful to track the effectiveness of performed control measures, which aims to push it below the epidemic threshold (corresponding to $R(t) = 1$). $R(t)$ incorporates factors affecting the spread of the epidemic (e.g., individual's behavior and susceptible depletion.

To estimate $R(t)$, we use the same methodology described previously (4–6) to distinguish between locally acquired and imported cases. Thus, we assume that the daily number of new cases (date of symptom onset) with locally acquired infection $L(t)$ can be approximated by a Poisson distribution

$$C(t)Pois(R_t \sum_{s=1}^{t} \phi(s)C(t-s)) \qquad [1]$$

where $C(t)$ is the number of new cases (either locally acquired or imported) at time t (date of symptom onset), $R(t)$ is the effective reproduction number at time t and is the generation time distribution. To estimate the time between consecutive generations of cases, we adopted the serial interval (which measures the time difference between the symptom onset of the infectors and of their infected) estimated from the literature (1), namely a gamma distribution with mean 5.0 days and standard deviation 3.4 days (shape=4.87, rate=0.65).

The likelihood λ of the observed time series cases from day 1 to T can be written as:

$$\lambda = \Pi_{t=1}^{T} P\left(C(t)Pois(R_t \sum_{s=1}^{t} \phi(s)C(t-s))\right) \qquad [2]$$

where $P(x, y)$ is the Poisson density distribution of observing x events, given the parameter y.

We then use Metropolis-Hastings MCMC sampling to estimate the posterior distribution of $R(t)$. The Markov chains were run for 100,000 iterations, considering a burn-in period of 10,000 steps, and assuming non-informative prior distributions of $R(t)$ (flat distribution in the range (0-1000]). Convergence was checked by visual inspection by running multiple chains starting from different starting points.

Constructing exposure measures. We constructed a measure of a colleges exposure to different geographic areas using movement data from 2020. For each county and college campus we estimated the average number of devices from the source county on each campus by week and computed the change from 20 to 14 days before reopening to 0 to 6 days after reopening. Because in some cases the net flow went in the opposite direction, we truncated the change in devices at 0.

We then estimated, for each county, 7-day incidence of COVID-19 ending 14 days before campus reopened. Finally, we used the truncated number of devices moving from each county to a given campus to construct the weighted average of all counties that had net movement towards a college campus. The resulting exposure metric is the average 7-day incidence for the period ending two weeks before campus reopened.

2. Robustness checks and alternative specifications

A. Trimming the sample. Our balanced panel includes all data for every county from June through November, which could include observations from as early as 14 weeks before or 13 weeks after reopening. We restricting our treated counties to being within five weeks of a reopening. Using this more restrictive, but unbalanced, sample yielded similar results as our main specification (row b Table S3), although our estimates for the changes in cases resulting in hospitalization and ICU admission are statistically significant and indicate that reopening a college campus increases the number of cases requiring hospitalization, ICU admission, or leading to mortality by 0.052, 0.015, and 0.045 cases per 100,000, respectively. In our restricted sample, we continue to find evidence that R_t was affected by campus reopenings. Figure S3 presents the corresponding event studies which are similar to the event studies for the balanced panel, although there may be more evidence of a differential time trend for hospitalization, ICU admissions, and mortality than in the balanced sample.

B. Restricting to counties with a college. Restricting our sample to counties with a college or university has minimal effects on our results–we continue to find significant increases in mobility and incidence, of roughly the same magnitude as our main results (row c, Table S3), although results for hospitalizations, ICU admissions, and mortality are estimated less precisely than in our main specification.

C. Weighting by county population. When we weight counties by population (row d, Table S3), in addition to weighting each college within a county by its enrollment share, our results vanish, which is what one should expect since in larger counties colleges and universities should have a smaller effect on COVID-19 incidence.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

D. Choosing one college per county. Assigning counties to treatment based on the college that reopened first (and largest, if there were ties) in each county yielded qualitatively similar results as our main specification (row e, Table S3), although we no longer find statistically significant evidence that reopening a college campus increases COVID-19 related hospitalizations.

E. Alternative time trends. To check if linear county-specific time trends adequately captured county deviations from national averages, We estimated our main difference-in-difference model including both linear and quadratic county-specific time trends (row f, Table S3), yielded consistent, but less precise, estimate as our main analysis. Omitting our linear time trends reversed our main findings for COVID-19 incidence and related outcomes (row g, Table S3). In some cases, the event studies (Figure S4) demonstrate that our negative coefficients are due to longer-term reductions in outcomes, relative to the control group. However, it is unclear if these reductions after 4, 5 , or more weeks reflect absolute reductions in our outcomes or an increase in the outcome in "control" counties.

F. Omitting date fixed effects. Our main models include date fixed effects. In row h of Table S3 we report results omitting these fixed effects in row . Omitting date fixed effects reverse our main findings, suggesting that reopening college campuses had a protective effect on an area. It is likely that this difference reflects the fact that most colleges reopened earlier in our sample period and the intensity of the COVID-19 pandemic rose throughout the Fall.

References

1. Du Z, et al. (2020) Serial interval of covid-19 among publicly reported confirmed cases. *Emerging infectious diseases* 26(6):1341.
2. Keeling MJ, Rohani P (2011) *Modeling Infectious Diseases in Humans and Animals*. (Princeton University Press).
3. Van Kerkhove MD, Bento AI, Mills HL, Ferguson NM, Donnelly CA (2015) A review of epidemiological parameters from ebola outbreaks to inform early public health decision-making. *Scientific Data* 2(1):150019.
4. Liu QH, et al. (2018) Measurability of the epidemic reproduction number in data-driven contact networks. *Proceedings of the National Academy of Sciences* 115(50):12680–12685.
5. Liu QH, et al. (2021) The covid-19 outbreak in sichuan, china: Epidemiology and impact of interventions. *PLOS Computational Biology* 16(12):1–14.
6. Cori A, Ferguson NM, Fraser C, Cauchemez S (2013) A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. *American Journal of Epidemiology* 178(9):1505–1512.

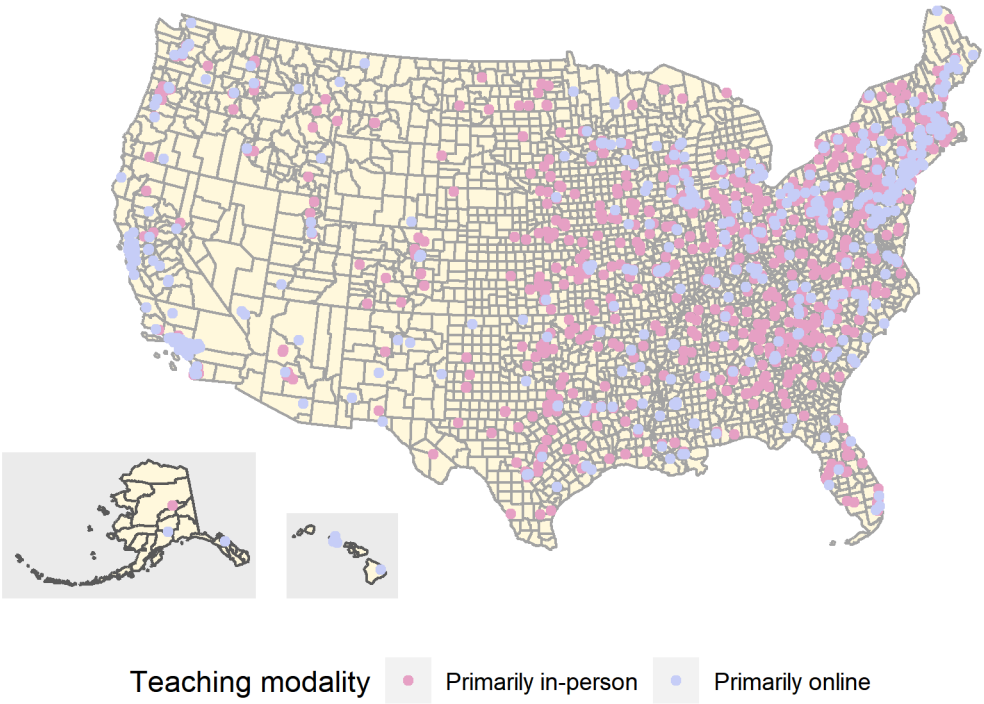


Fig. S1. Geographic distribution of colleges and universities in sample by teaching modality. Colleges and universities are more prevalent in the eastern half of the United States, while colleges in the western half were more likely to reopen for online teaching.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

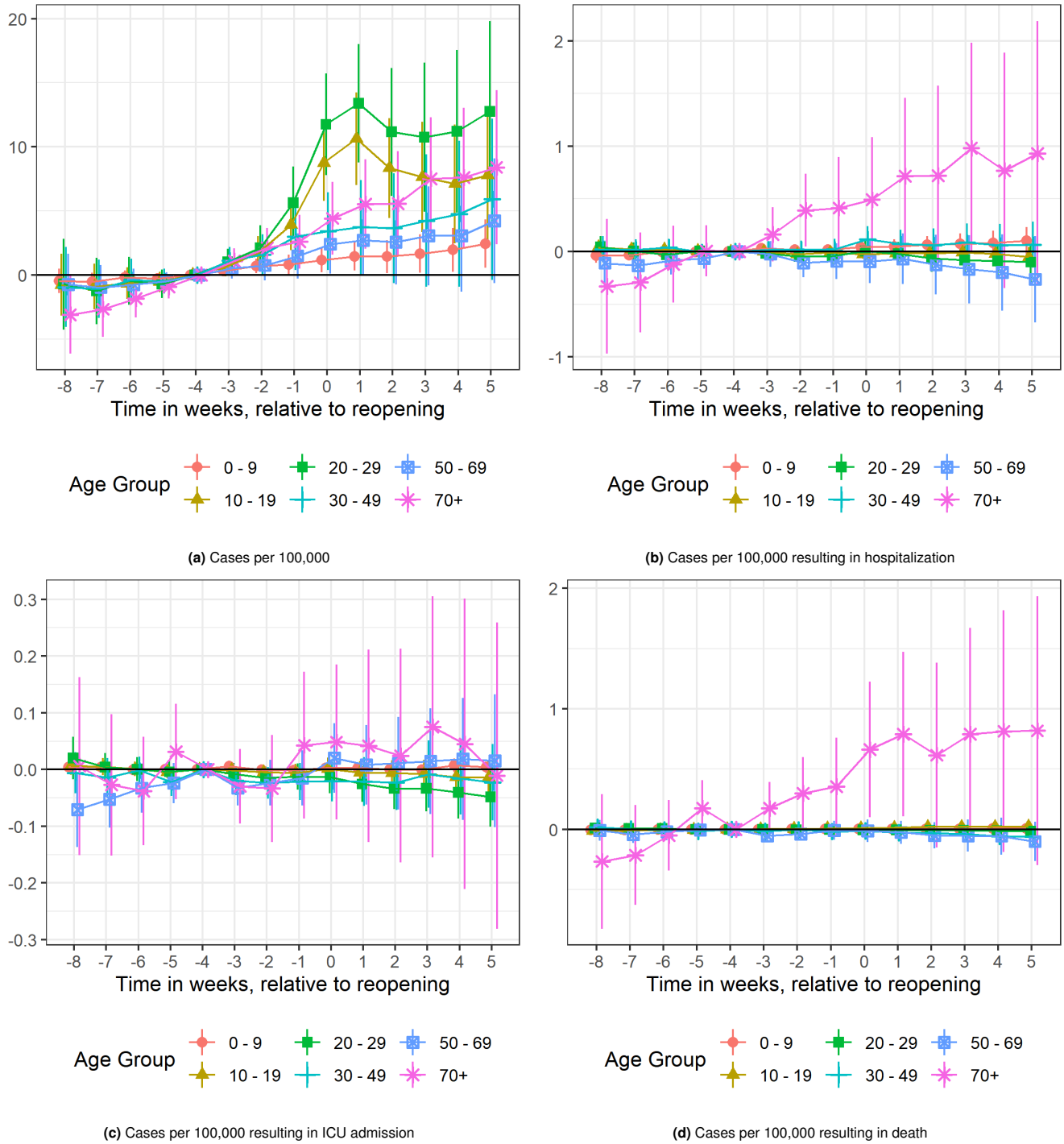


Fig. S2. Age-specific event studies. COVID-19 incidence rose for all age groups, following the resumption of classes, although the increase faded over time for some groups (a), while cases requiring hospitalization (b) and ICU care (c) followed less precisely estimated patterns. Mortality due to COVID-19 rose sharply for people 70 and older around the return of students and resumption of classes (d).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

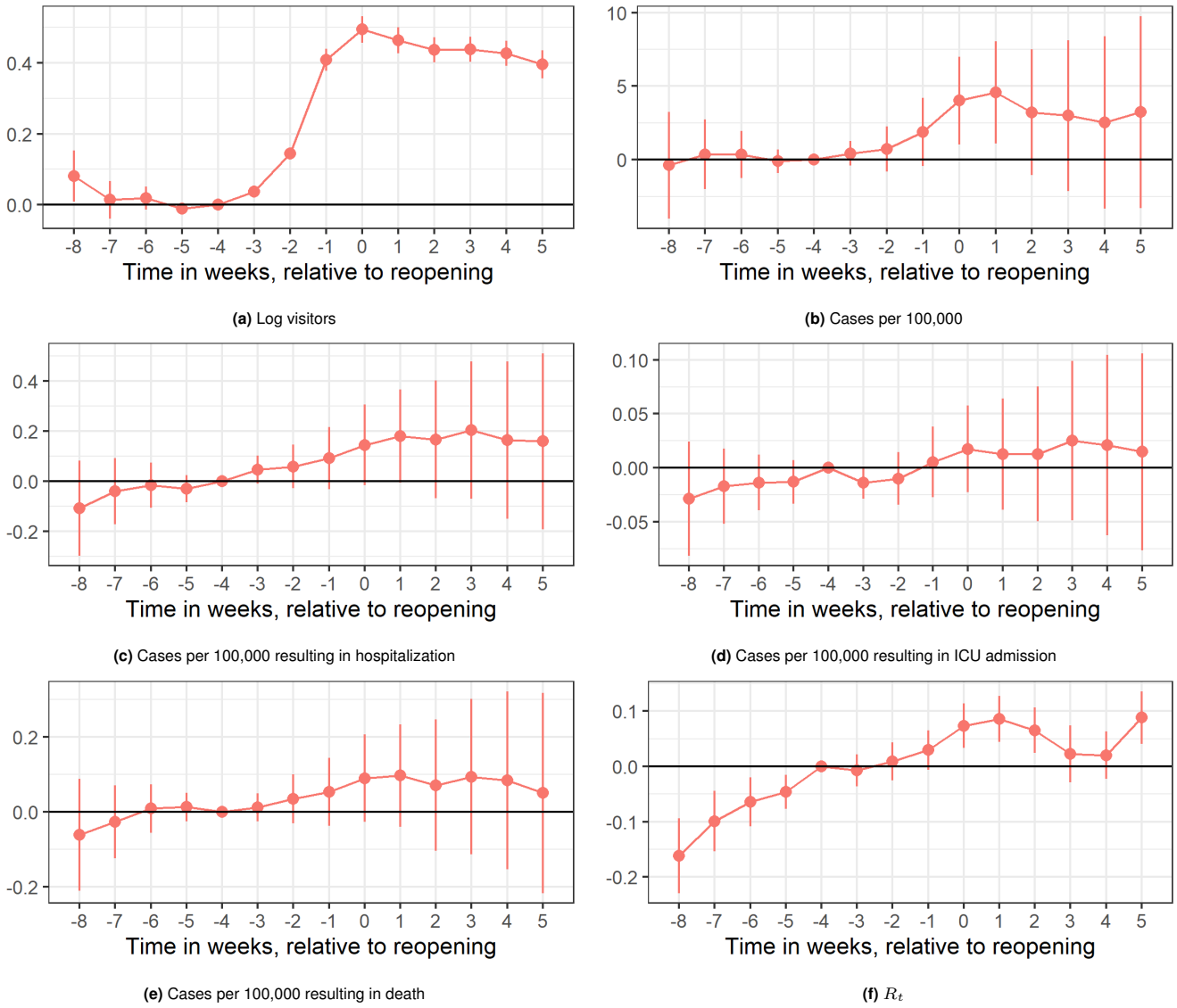
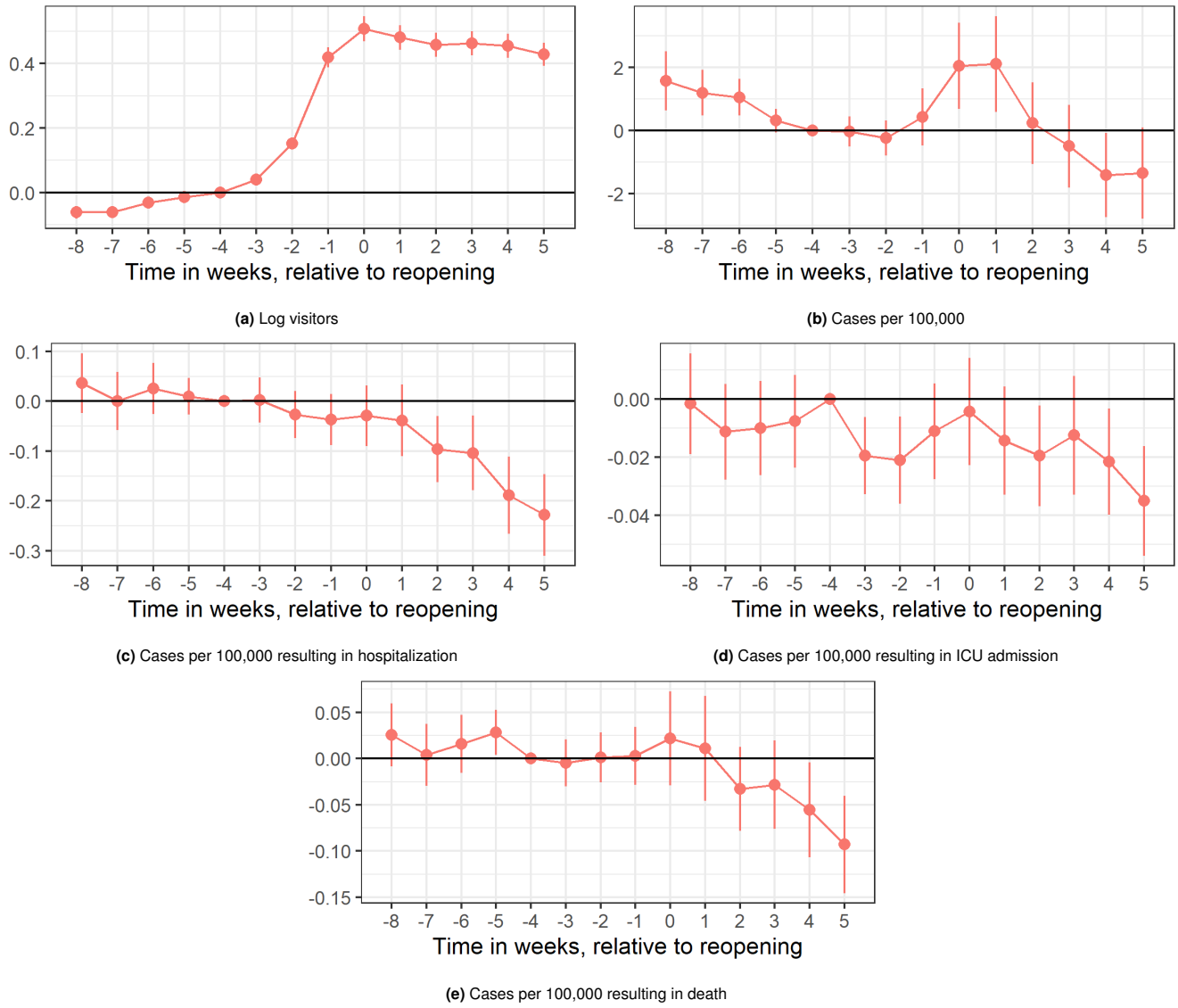


Fig. S3. Event studies using "trimmed" sample. Mobility increased in the lead up to reopening and remained high following the resumption of classes (a). Increases in cases followed the rise in mobility by one week (b), with a rising trend in cases resulting hospitalization (c) and deaths (e), but not ICU admissions (d). R_t rose following reopening as well, following the rise in mobility by about one week (f).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



(a) Log visitors

(b) Cases per 100,000

(c) Cases per 100,000 resulting in hospitalization

(d) Cases per 100,000 resulting in ICU admission

(e) Cases per 100,000 resulting in death

Fig. S4. Event studies without time trends. Models without time trends continue to demonstrate a substantial increase in on-campus mobility (a). However, the post-reopening increase in COVID-19 incidence is transient (b). Results for hospitalizations (c), ICU admissions (d), and mortality (e) exhibit downward trends following the resumption of classes.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table S1. Difference-in-difference regressions demonstrating the effect of reopening college campuses on mobility and COVID-19 incidence and sequelae.

	(1) Log visitors	(2) Daily new cases per 100,000	(3) Daily new hospitalized cases per 100,000	(4) Daily new cases in ICU per 100,000	(5) Daily new deaths per 100,000	(6) Rt
Baseline model						
Post-reopening	0.413***	2.728***	0.052+	0.013+	0.042*	0.164***
	(0.016)	(0.667)	(0.031)	(0.008)	(0.020)	(0.014)
By teaching modality						
Primarily in-person	0.471***	3.190***	0.063+	0.013+	0.051*	0.160***
	(0.019)	(0.752)	(0.033)	(0.008)	(0.021)	(0.015)
Primarily Online	0.330***	1.698*	0.028	0.013	0.022	0.173***
	(0.024)	(0.835)	(0.041)	(0.012)	(0.029)	(0.017)
Prior and post reopening						
1 week before reopening	0.414***	1.371**	0.012	0.008	0.019	0.099***
	(0.016)	(0.524)	(0.027)	(0.007)	(0.016)	(0.016)
0-1 weeks after reopening	0.487***	3.721***	0.042	0.013	0.045+	0.154***
	(0.019)	(0.796)	(0.036)	(0.009)	(0.026)	(0.018)
2-3 weeks after reopening	0.449***	2.480**	0.011	0.010	0.014	0.125***
	(0.018)	(0.901)	(0.047)	(0.011)	(0.031)	(0.019)
4+ weeks after reopening	0.379***	2.470*	-0.068	0.001	-0.017	0.191***
	(0.016)	(0.962)	(0.055)	(0.013)	(0.039)	(0.016)
Interacted with exposure						
Post-reopening	0.409***	2.534***	0.058+	0.013+	0.042*	0.164***
	(0.015)	(0.653)	(0.031)	(0.008)	(0.020)	(0.014)
x Exposure (x 10,000)	0.157***	1.611*	-0.047+	-0.001	0.002	-0.001
	(0.024)	(0.708)	(0.027)	(0.006)	(0.017)	(0.006)

Source—Authors' analysis of C2I data on college reopening, SafeGraph mobility data, and COVID-19 case and mortality data.

Notes—All models include county, day, and college fixed effects, and county-specific time trends (except for R_t); each college within a county is weighted by its share of total enrollment in the county. "Primarily in-person" includes "Fully in-person" and "Hybrid"; "Primarily online" includes "Fully online". Exposure is the 7-day COVID-19 incidence in source counties for each college two weeks prior to reopening, weighted by the change in movement to the campus from each county. Column titles indicate the dependent variable; each panel is a separate specification. Standard errors clustered on county in parentheses. + P<0.1, * P<0.05, ** P<0.01, *** P<0.001

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table S2. Age-specific difference-in-differences demonstrates that the increase in incidence of COVID-19 was isolated to people between 10 and 29 years of age, which encompasses college-age individuals

Age group	(1) Daily new cases per 100,000	(2) Daily new hospitalized cases per 100,000	(3) Daily new cases in ICU per 100,000	(4) Daily new deaths per 100,000
0 - 9	-0.137	-0.007	-0.000	-0.003+
	(0.237)	(0.011)	(0.002)	(0.001)
10 - 19	6.610***	0.017	0.003	0.002
	(0.961)	(0.012)	(0.003)	(0.002)
20 - 29	6.994***	0.031	0.005	-0.001
	(1.168)	(0.021)	(0.005)	(0.003)
30 - 39	-0.041	0.011	0.000	0.002
	(0.895)	(0.031)	(0.009)	(0.007)
40 - 49	0.096	0.000	-0.000	0.003
	(0.814)	(0.045)	(0.011)	(0.016)
50 - 59	0.227	-0.050	-0.001	0.018
	(0.642)	(0.048)	(0.015)	(0.017)
60 - 69	0.667	0.022	0.010	0.043
	(0.592)	(0.077)	(0.023)	(0.042)
70+	0.825	0.085	0.083*	0.127
	(0.705)	(0.146)	(0.036)	(0.135)

Source—Authors' analysis of C2I data, SafeGraph mobility data, and CDC COVID-19 case data.

Notes—All models include county, day, and college fixed effects, and county-specific time trends; each college within a county is weighted by its share of total enrollment in the county. Column titles indicate the dependent variable; models estimated separately by age group. Standard errors clustered on county in parentheses. + P<0.1, * P<0.05, ** P<0.01, *** P<0.001

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table S3. Robustness checks discussed in section 2

	Log visitors	Daily new cases per 100,000	Daily new hospitalized cases per 100,000	Daily new cases in ICU per 100,000	Daily new deaths per 100,000	Rt
a) Base line	0.413***	2.728***	0.052+	0.013+	0.042*	0.164***
	(0.016)	(0.667)	(0.031)	(0.008)	(0.020)	(0.014)
b) Restrict to ±5weeks[a]	0.397***	2.701***	0.052*	0.015*	0.045**	0.066***
	(0.016)	(0.506)	(0.024)	(0.006)	(0.017)	(0.013)
c) Counties with colleges	0.412***	2.629***	0.032	0.009	0.030+	0.014
	(0.016)	(0.484)	(0.023)	(0.006)	(0.018)	(0.014)
d) Weight by population	0.283***	1.529+	-0.045	-0.000	0.021	0.160***
	(0.036)	(0.913)	(0.049)	(0.004)	(0.016)	(0.009)
e) First and largest college		2.319***	0.050+	0.012+	0.044*	0.165***
		(0.637)	(0.030)	(0.007)	(0.019)	(0.014)
f) Quadratic time trends	0.409***	2.382***	0.015	0.011	0.023	-0.019
	(0.016)	(0.496)	(0.024)	(0.007)	(0.018)	(0.018)
g) Drop linear time trends	0.431***	-1.625**	-0.182***	-0.014*	-0.070***	
	(0.017)	(0.621)	(0.031)	(0.005)	(0.017)	
h) Drop date fixed effects	0.436***	7.198***	0.083***	-0.001	0.073***	-0.083***
	(0.017)	(0.527)	(0.022)	(0.004)	(0.012)	(0.008)

[a] Sample restricted to ± 5 weeks includes all counties without a college, but only includes observations for each county with a college that is within five weeks of a college reopening.

Source–Authors' analysis of C2I data, SafeGraph mobility data, and CDC COVID-19 case data.

Notes–Each panel presents results for a separate set of regression models. Each column corresponds to a different sample and/or weighting approach. Baseline model includes county, day, and college fixed effects, and county-specific time trends (except for R_t). Standard errors clustered on county in parentheses.

* P<0.05, ** P<0.01, *** P<0.001

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

……PATIENT IDENTIFIER INFORMATION IS NOT TRANSMITTED TO CDC……

87

Patient first name _____ Patient last name _____ Date of birth (MM/DD/YYYY): ____/_____/_____

……PATIENT IDENTIFIER INFORMATION IS NOT TRANSMITTED TO CDC……

Human Infection with 2019 Novel Coronavirus Case Report Form

Reporting Jurisdiction		Case state/local ID	
Reporting Health Department		CDC 2019-nCoV ID	
Contact ID[a]		NNDSS loc. rec. ID/Case ID[b]	

[a]Only complete if case-patient is a known contact of prior source case-patient. Assign Contact ID using CDC 2019-nCoV ID and sequential contact ID, e.g., Confirmed case CA102034567 has contacts CA102034567 -01 and CA102034567 -02. [b]For NNDSS reporters, use GenV2 or NETSS patient identifier.

Interviewer Information

Name of Interviewer: Last: First: Telephone: Email:

Affiliation/Organization:

Case Classification and Identification

What is the current status of this person?

☐ Lab-confirmed case* ☐ Probable case

If probable, select reason for case classification:

☐ Meets clinical criteria AND epidemiologic evidence with no confirmatory lab testing*
☐ Meets presumptive lab evidence± AND either clinical criteria OR epidemiologic evidence
☐ Meets vital records criteria with no confirmatory lab testing

*Detection of SARS-CoV-2 RNA in a clinical specimen using a molecular amplification detection test

± Detection of specific antigen in a clinical specimen, OR detection of specific antibody in serum, plasma, or whole blood indicative of a new or recent infection

Under what process was the case first identified? (check all that apply)

☐ Clinical evaluation ☐ Routine surveillance
☐ Contact tracing of case patient ☐ Other, specify: _____
☐ EpiX notification of travelers. If yes, DGMQID: _____
☐ Unknown

Report date of case to CDC (MM/DD/YYYY):
___/_____/_____

Date of first positive specimen collection (MM/DD/YYYY):
___/_____/_____ ☐ Unknown ☐ N/A

Hospitalization, ICU, and Death Information

Was the patient hospitalized?
☐ Yes ☐ No ☐ Unknown

If yes, admission date 1 ___/___/_____ (MM/DD/YYYY) discharge date 1 __/___/_____

If hospitalized, was a translator required?
☐ Yes ☐ No ☐ Unknown

If yes, specify which language: _____

Was the patient admitted to an intensive care unit (ICU)?
☐ Yes ☐ No ☐ Unknown

If yes, admission date 1 ___/___/_____ (MM/DD/YYYY) discharge date 1 __/___/_____

Did the patient die as a result of this illness?
☐ Yes ☐ No ☐ Unknown If yes, date of death (MM/DD/YYYY): ____/_____/_____ ☐ Unknown date

Case Demographics

Date of birth (MM/DD/YYYY): ____/_____/_____
Age: _____ Age units (yr/mo/day): _____
State of residence: ____ County of residence: _____
Does this case have any tribal affiliation? ☐ yes
Tribe name(s): _____ Enrolled member? ☐ yes

Sex:
☐ Male ☐ Other
☐ Female ☐ Unknown

If female, currently pregnant?
☐ Yes ☐ No ☐ Unknown

Ethnicity:
☐ Hispanic/Latino
☐ Non-Hispanic/Latino
☐ Unknown

Race (check all that apply):
☐ Black ☐ White ☐ Asian
☐ American Indian/Alaska Native
☐ Native Hawaiian/Other Pacific Islander
☐ Unknown ☐ Other, specify: _____

Which would best describe where the patient was staying at the time of illness onset?

☐ House/single family home ☐ Hotel/motel ☐ Nursing home/assisted living facility ☐ Rehabilitation facility ☐ Mobile home
☐ Apartment ☐ Long term care facility ☐ Acute care inpatient facility ☐ Correctional facility ☐ Group home
☐ Homeless shelter ☐ Outside, in a car, or other location not meant for human habitation ☐ Other (specify): _____ ☐ Unknown

Healthcare Worker Information

Is the patient a health care worker in the United States? ☐ Yes ☐ No ☐ Unknown

If yes, what is their occupation (type of job)?
☐ Physician ☐ Respiratory therapist ☐ Other, specify: _____
☐ Nurse ☐ Environmental services ☐ Unknown

If yes, what is their job setting?
☐ Hospital ☐ Rehabilitation facility ☐ Other, specify:_____
☐ Long-term care facility ☐ Nursing home/assisted living facility ☐ Unknown

Exposure Information

In the 14 days prior to illness onset, did the patient have any of the following exposures (check all that apply):

☐ Domestic travel (outside state of normal residence). Specify state(s): _____
☐ International travel. Specify country(s): _____
☐ Cruise ship or vessel travel as passenger or crew member. Specify name of ship: ____
☐ Workplace
If yes, is the workplace critical infrastructure (e.g., healthcare setting, grocery store)?
☐ Yes, specify workplace setting: _____ ☐ No ☐ Unknown
☐ Airport/airplane
☐ Adult congregate living facility (nursing, assisted living, or long-term care facility)
☐ School/university/childcare center
☐ Correctional facility
☐ Community event/mass gathering
☐ Animal with confirmed or suspected COVID-19. Specify animal: _____
☐ Other exposures, specify: _____
☐ Unknown exposures in the 14 days prior to illness onset

☐ Contact with a known COVID-19 case (probable or confirmed)

If the patient had contact with a known COVID-19 case:
What type of contact?
☐ Household contact
☐ Community-associated contact
☐ Healthcare-associated contact (patient, visitor, or healthcare worker)

Was this person a U.S. case?
☐ Yes, nCoV ID(s) _____, _____, _____
☐ No, this person was an international case and contact occurred abroad
☐ Unknown if U.S. or international case

Is this case part of an outbreak?
☐ Yes, specify outbreak name: _____ ☐ No ☐ Unknown

medRxiv preprint doi: https://doi.org/10.1101/2020.09.22.20196048; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

……PATIENT IDENTIFIER INFORMATION IS NOT TRANSMITTED TO CDC……

Patient first name _____ Patient last name _____ Date of birth (MM/DD/YYYY): ____/_____/_____

……PATIENT IDENTIFIER INFORMATION IS NOT TRANSMITTED TO CDC……

Human Infection with 2019 Novel Coronavirus
Case Report Form

Clinical course, symptoms, past medical history, and social history

Collected from (check all that apply): ☐ Patient interview ☐ Medical record review

Symptoms present during course of illness:

☐ Symptomatic
☐ Asymptomatic
☐ Unknown

If case was symptomatic:

What was the onset date?
Onset date (MM/DD/YYYY): __/__/____
☐ Unknown symptom onset date

Did the patient's symptoms resolve?
Date of symptom resolution (MM/DD/YYYY): __/__/_____
☐ No, still symptomatic
☐ Symptoms resolved, unknown date
☐ Unknown if symptoms resolved

Did the patient develop pneumonia?
☐ Yes ☐ No ☐ Unknown

Did the patient have acute respiratory distress syndrome?
☐ Yes ☐ No ☐ Unknown

Did the patient have an abnormal chest X-ray?
☐ Yes ☐ No ☐ Unknown ☐ N/A, no chest X-ray done

Did the patient have another diagnosis/etiology for their illness?
☐ Yes ☐ No ☐ Unknown

Did the patient have an abnormal EKG?
☐ Yes ☐ No ☐ Unknown ☐ N/A, no EKG done

Did the patient receive mechanical ventilation (MV)/intubation?
☐ Yes ☐ No ☐ Unknown
If yes, total days with MV (days) _____

Did the patient receive ECMO?
☐ Yes ☐ No ☐ Unknown

If symptomatic, which of the following did the patient experience during their illness?

Fever >100.4F (38C)[c]	☐Yes	☐No	☐Unk	Cough (new onset or worsening of chronic cough)	☐Yes	☐No	☐Unk
Subjective fever (felt feverish)	☐Yes	☐No	☐Unk	Wheezing	☐Yes	☐No	☐Unk
Chills	☐Yes	☐No	☐Unk	Shortness of breath (dyspnea)	☐Yes	☐No	☐Unk
Rigors	☐Yes	☐No	☐Unk	Difficulty breathing	☐Yes	☐No	☐Unk
Muscle aches (myalgia)	☐Yes	☐No	☐Unk	Chest pain	☐Yes	☐No	☐Unk
Runny nose (rhinorrhea)	☐Yes	☐No	☐Unk	Nausea or vomiting	☐Yes	☐No	☐Unk
Sore throat	☐Yes	☐No	☐Unk	Abdominal pain	☐Yes	☐No	☐Unk
New olfactory and taste disorder(s)	☐Yes	☐No	☐Unk	Diarrhea (≥3 loose stools/24hr period)	☐Yes	☐No	☐Unk
Headache	☐Yes	☐No	☐Unk	Other, specify: _____, _____, _____	☐Yes	☐No	☐Unk
Fatigue	☐Yes	☐No	☐Unk				

Did they have any underlying medical conditions and/or risk behaviors? ☐Yes ☐No ☐ Unknown

Diabetes Mellitus	☐Yes	☐No	☐Unk	Immunosuppressive condition	☐Yes	☐No	☐Unk
Hypertension	☐Yes	☐No	☐Unk	Autoimmune condition	☐Yes	☐No	☐Unk
Severe obesity (BMI ≥40)	☐Yes	☐No	☐Unk	Current smoker	☐Yes	☐No	☐Unk
Cardiovascular disease	☐Yes	☐No	☐Unk	Former smoker	☐Yes	☐No	☐Unk
Chronic Renal disease	☐Yes	☐No	☐Unk	Substance abuse or misuse	☐Yes	☐No	☐Unk
Chronic Liver disease	☐Yes	☐No	☐Unk	Disability (neurologic, neurodevelopmental, intellectual, physical, vision or hearing impairment) If yes, specify: _____	☐Yes	☐No	☐Unk
Chronic Lung disease (asthma/emphysema/COPD)	☐Yes	☐No	☐Unk				
Other chronic diseases If yes, specify: _____	☐Yes	☐No	☐Unk				
Other underlying condition or risk behavior, specify: _____	☐Yes	☐No	☐Unk	Psychological/psychiatric condition If yes, specify: _____	☐Yes	☐No	☐Unk

SARS-CoV-2 Testing (approved by FDA or other designated authority)

Test	Pos	Neg	Indet./Inconc.	Pend.	Not Done
Molecular amplification test (RT PCR)	☐	☐	☐	☐	☐
Serologic test	☐	☐	☐	☐	☐
Other (specify): _____	☐	☐	☐	☐	☐

Specimens for CoV-19 Testing

Specimen ID
1) _____
2) _____
3) _____

Additional Comments or Notes