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

Martin S. Andersen[1], Ana I. Bento[2], Anirban Basu[3], Chris Marsicano[4], Kosali Simon[5]

1 Department of Economics, University of North Carolina at Greensboro, Greensboro, NC 27412
2 Department of Epidemiology and Biostatistics, School of Public Health, Indiana University-Bloomington, Bloomington, IN 47405
3 The Comparative Health Outcomes, Policy, and Economics (CHOICE) Institute, Departments of Pharmacy, Health Services, and Economics, University of Washington, Seattle, Seattle, WA 98195
4 The College Crisis Initiative at Davidson College, Assistant Professor of the Practice of Higher Education Department of Educational Studies, Davidson College, Davidson, North Carolina 28035
5 O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington, IN 47405

Corresponding Author:
Martin S. Andersen
Department of Economics
University of North Carolina at Greensboro
516 Stirling Street,
Greensboro NC, 27412
Email: Msander4@uncg.edu

Competing Interest Statement: Dr. Marsicano's was supported in part by a grant from the ECMC Foundation.

Classification: Social Sciences – Economic Sciences

Keywords: COVID-19; college reopening; disease transmission

Significance Statement

The effect of reopening college campuses on COVID-19 cases is unknown. We demonstrate that reopening college campuses is associated with increases in movement on those campuses and an increase in COVID-19 incidence in the county. The increase in COVID-19 cases is larger for counties that contain a campus that reopened for in-person instruction. College administrators and others should continue the potential impact of reopening on COVID-19 cases as they make plans for the Spring semester.

Abstract

Most U.S. colleges have reopened campuses for in-person teaching this Fall, following rapid closures at the onset of the COVID-19 pandemic this Spring. Despite administrative efforts at mitigation and preventive measures, the large congregation of students within close quarters has caused public health concerns. In this paper, we examine college reopenings' association with changes in human mobility within campuses and in COVID-19 incidence in the counties of the campuses, over a two-week period before and after college reopenings. To estimate the daily reproduction number (Rt), we used a Bayesian framework. Using a difference-in-differences design comparing areas with a college campus, before and after reopening, to areas without a campus, we find that after college reopenings for face-to-face instruction, COVID-19 incidence in the county increased on average by a statistically significant 0.024 per thousand residents, following increases in mobility on campus. Similarly, we estimated increased transmission locally after reopening, with a rising trend in R_t. The increase in cases was larger in counties with colleges that drew students from areas with increasing incidence rates.

During late Summer 2020, colleges reopened and welcomed hundreds of thousands of students back to campus in the United States (1). However, in several prominent cases, institutions switched to online instruction after rapid increases in COVID-19 cases on campus (2). These early experiences call into question the feasibility of resuming in-person instruction during the COVID-19 pandemic. Several reports indicate that opening may have promoted the spread of infections among students, and likely in the community (3). However, we are unaware of any research that formally tests this hypothesis in the U.S. or elsewhere. We hypothesized that reopening colleges would increase COVID-19 cases within the college community with potential spillovers onto the neighboring areas. We also hypothesized that increases in incidence would be greater in campus' that attract students from areas with higher incidence of COVID-19, and that effects would be concentrated among campuses providing face-to-face instruction. 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 more recently.

We describe our methods in the Supplementary Information.

Results

Our sample period ran from July 15th 2020 to September 13th 2020. Across the 3,142 counties of the United States, only 779 contained a college campus in our universe of 1,409 colleges (Panel A of **eFigure 1**).

The number of devices visiting campus increased significantly, in the week leading up to the start of classes and after classes have begun (Panel B, **Figure 1**). In the aggregate, the number of devices on campus rose by 47.3% (**eTable 1** Model 1, 95% CI: 36.3% - 58.2%, (4)) in the 2 weeks after classes started, relative to the 2 week period before the start of class. In **eTable 2**, we break out these results further by week relative to the reopening date. The increase was larger among schools that opened for primarily in-person education (55.7%; CI: 43.0 - 68.5) versus those that did not (33.2%, CI: 23.4% - 43.0%) (**eTable 1** Model 2, (4)). **eFigures 1** and **2** (4) present conventional event studies that use counties with a school that has not yet opened, or without a school, as a comparison group and yield similar conclusions.

Using our difference-in-difference framework, we find that reopening a college was associated with a statistically significant increase of 0.017 cases per 1000 people (**eTable 1**, Model 1, (4)). 95% CI: 0.003-0.030). When we separate the types of opening (Model 2), we find that reopening for primarily in-person instruction increases COVID-19 cases in the county by 0.024 (Model 2, 95% CI: 0.008 - 0.040) cases per thousand, with no significant effect of opening online. There trends are illustrated in Figure 2.

Using our weighted measure created for exposure to cases from student –out-of-county migration, we find that a campus with a ten percent higher exposure level greater exposure is associated with an additional 0.0119 cases per thousand county residents (Model 3, 95% CI: 0.0078 – 0.0160) per day than other campuses, holding all else constant. Further, when including in-person instruction interactions, we find that the ten percent increase in the out-of-county exposure measure is associated with an

additional 0.0142 (CI: 0.0092 - 0.0192) cases per thousand in counties in which the first school opened for in-person instruction and no association for other counties.

We performed several additional analyses and specification checks, which we report in GitHub and describe here. In **eTables 3-5** (4), we report similar results using alternative measures of our main disease-related outcome measure. In **eTable 3**, (4) we use the number of new cases over the last three days, which ensures that results are not driven by weekends. **eTable 4** (and **eFigure 1**, Panel C) use new cases over a week, and **eTable 5** (and **eFigure 1**, Panel D) report results using R_t as the dependent variable. Using the R_t estimates, the main effect of reopening is consistently significant and indicates that reopening was associated with an increase transmission, we no longer find significant differences by instructional modality**.** Weighting counties by population (**eTable 6**, (4)), resulted in no significant differences between online and in-person reopening, nor was there a statistically significant association with reopening. However, we continued to find that reopening schools that drew students from higher risk areas was associated with an increase in COVID-19 cases in the county. In addition, we used higher resolution measures of in-person teaching (**eTable 7**, (4)), which indicate that there is a dose-response relationship between the degree of in-person teaching and new cases. Institutions that opened for primarily in-person instruction and those that adopted hybrid or hyflex models were predicted to have higher numbers of visitors than those at the extremes of the in-person to online distribution.

Discussion

Our findings demonstrate that re-opening a college is associated with an increase in the number of cellular devices in the week leading up to the start of classes, irrespective of type of instruction offered. Our second main finding is that re-opening a college with in-person instruction significantly increases the rate of new daily cases in the county. We also find a significant and positive relationship between daily new cases in the county and a county's exposure to new cases from students' home states in the two weeks before reopening. This is consistent with the county estimates of Rt, reflecting a steady increase in transmission 2weeks after the uptick in mobility.

Using our classification of counties, our results indicate that reopening college campuses for in-person instruction is associated with more than 3000 (3219 [95% CI: 1094 – 5344]) additional cases of Covid-19 per day in the United States , assuming that our results are not due solely to new cases being diagnosed on campus, rather than at home. 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 to this, asymptomatic cases will only be caught if testing was done in 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 in the first week of reopening, assuming that a positive test result is returned nine days after infection, while we find no increase in COVID-19 cases until the second week of classes, as reflected by Rt.

We did not quantify potential spillovers to the communities surrounding campuses, as these effects will take more weeks to be observed and would benefit from college-level incidence data. 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 also beyond the scope of this initial study—these remain priorities in forthcoming studies. Several other limitations of our analysis arising primarily out of lack of data are discussed in the SI.

We submit, our findings are critical in the context of public health adaptive management strategies and in particular for colleges, as they consider additional strategies to mitigate disease burden and decrease transmission. Institutional leaders should think carefully as they plan their Spring 2020 semesters – examining not only the "situation on the ground" in their own institutions' communities, but also in those from which they draw a substantial number of students.

Acknowledgements
SafeGraph, a data company that aggregates anonymized location data from numerous applications in order to provide insights about physical places, provided data for this project. To enhance privacy, SafeGraph excludes census block group information if fewer than five devices visited an establishment in a month from a given census block group.

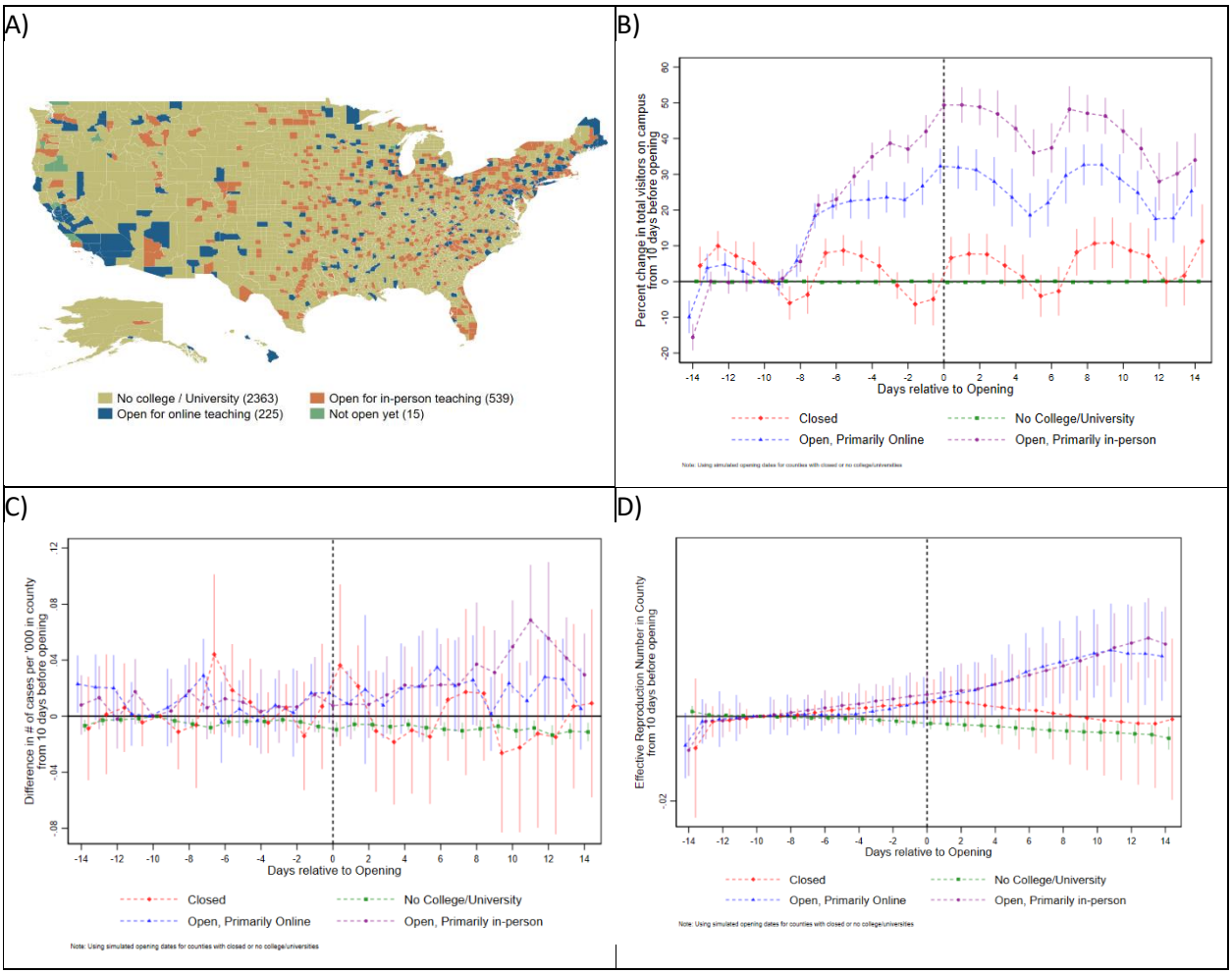


Figure1: Classification of counties (A) and event Studies of Mobility, Cases, and R_t around reopening by reopening status of the Census Block Group or County. Most schools began allowing students to move-in around seven days before the start of class (B). Cases in counties with a school that opened for in-person education increased significantly by 8 days after reopening (C). Trends in R_t (D) indicate that reopening, rather than instructional modality, were associated with changes in R_t over time. Each panel reports estimates from a single regression model including controls for county/CBG and date. 95% confidence intervals based on standard errors clustered on county and allowing for serial correlation in the error term.

References

1. Marsicano C. C2I: The College Crisis Initiative – Crisis to Innovation [Internet]. [cited 2020 Sep 14]. Available from: https://collegecrisis.org/

2. Nierenberg A, Pasick A. Schools Briefing: University Outbreaks and Parental Angst. The New York Times [Internet]. 2020 Aug 19 [cited 2020 Sep 14]; Available from: https://www.nytimes.com/2020/08/19/us/colleges-closing-covid.html

3. Hubler S, Hartocollis A. How Colleges Became the New Covid Hot Spots. The New York Times [Internet]. 2020 Sep 11 [cited 2020 Sep 14]; Available from: https://www.nytimes.com/2020/09/11/us/college-campus-outbreak-covid.html

4. Andersen M, Bento AI, Basu A, Marsicano C, Simon K. anabento/COVID19-CollegeReopening [Internet]. 2020 [cited 2020 Sep 20]. Available from: https://github.com/anabento/COVID19-CollegeReopening

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

Supplemental Information

A. Methods

A1. Study Data

We collected data on opening dates and announced instructional methods from the College Crisis Initiative(1) 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 merged college opening dates with "shapefiles" (the geographic coordinates) for college campuses (2) which allowed us to assign college campuses to Census Block Groups (CBGs) and counties. Our final sample included 1,404 colleges, as 27 colleges did not map to a CBG. We dropped an additional 2 schools that did not have an announced reopening strategy for a final sample of 1,409 schools. 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 (886) or primarily online (483). Across these colleges there are campuses in 1446 college-county pairs.

We extracted cellular data from SafeGraph's Social Distancing Metrics datafiles, an opt-in sample of approximately 20 million mobile devices. These data measure the number of devices that are detected each day in each CBG, from July 15th through Sunday September 13th. SafeGraph data have been used in several recent publications (3–5).

We collected daily state and county confirmed COVID-19 cases for the same period (6) dividing by U.S. Census population estimates to form our main outcome. In addition to case counts, we estimated the county-level net reproduction number (R_t) (7) (details in the eMethods) as a measure of disease transmission.

We used data from 2019 to construct estimates of each college's exposure to counties around the United States (see supplement for details) based on the number of devices moving onto campus around the start of classes for the Fall 2019 semester. We use 2019 data because that allows us to create potential exposure indices for all campuses, including those who have not reopened. Using our estimate of a college's exposure to counties, we computed the weighted average of the log weekly incidence just prior to campus reopening in 2020, for those areas from which a college drew students. We used the

change in this measure from July 15th to two weeks before reopening as a measure of the unexpected incidence in students' home areas.

A2. Effective reproductive number (R_t)

We used the daily cases, incubation period, and serial interval previously estimated (7). This allowed us to estimate the effective reproduction number for each county. The basic reproduction number R0 characterizes the mean number of secondary cases produced by a primary infector during the exponential growth phase of the epidemic, before interventions are applied and when the depletion of susceptible individuals is negligible (8). The effective reproduction number R(t) represents the mean number of secondary cases generated by a primary infector at time t (9). 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 (9,10) 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

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

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 (7), 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:

$$\Lambda = \prod_{t=1}^{T} P(C(t) \sum_{s=1}^{t} \phi(s)C(t-s)\ ),$$

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.

A3. Constructing exposure measures

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

We constructed a measure of a colleges exposure to different geographic areas using movement data from 2019. We collected start dates for the Fall 2019 semester and estimated the change in the number of devices on campus from a pre-period, which we defined as the period from 14 to 8 days before the start of classes, to a post-period, which we defined as the first week of classes. Because the number of devices in each original county varied over time, we normalized the visitors from each county by the ratio of the average number of devices from July 15 2019 through September 30 2019 to the number of devices observed on day t. We omitted movements that resulted in a decrease in devices on campus and then converted these movement changes into a measure of exposure as the fraction of devices moving to campus from county c relative to the total number of devices moving to campus.

We converted our empirical measure of movement into a measure of disease exposure by averaging the log number of new cases per 1000 over the week ending two weeks before the start of classes. Lastly, because it is possible that campus administrators considered the disease exposure of their students when choosing to reopen for face-to-face instruction, we computed the residual from a regression of our disease exposure measure from fourteen days before campus opened on disease exposure as of July 15th.

A4. Statistical Analysis

Our main analyses use a balanced panel of counties and CBGs. In cases where a county or CBG has multiple colleges represented (as happens for 266 of 3139 counties and 110 of 216992 CBGs), we only include the college with the earliest start date and broke ties by choosing the college with the highest total enrollment.

We used event-study and difference-in-difference methods to assess the relationship between college reopenings and our two main outcomes: mobility to campuses, and county daily COVID-19 reported cases.

Our event study assessed the changes in mobility and changes in COVID-19 cases, 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 cases models)—calendar date, and college (since a college can span more than one CBG or county). We constructed counterfactual time trends for counties without a college or that had not yet opened using the distribution of opening dates for all other college in our data-e.g. since 11.8% opened on Aug 31st, the counterfactual will assume that 11.8% of nonopened colleges opened on that date.

Our regression models for the event studies, using i to denote the geographic unit (CBG or county), can be written as:

$$O_{it} = \sum_{\substack{\tau=-14 \\ \tau\neq-10}}^{14} \beta_\tau \mathbf{1}[t - R_i = \tau] + \delta_i + sigma_t + \epsilon_{it}$$

Where $O_{it}$is the outcome, R_i is the reopening date for campus i, δ_i and σ_t are date and time fixed effects, and ϵ_{it} is an idiosyncratic error term. In some cases, we computed these event studies allowing

for different daily coefficients for groups of counties or CBGs defined by the opening status of the county (no college, open online, open face-to-face, not yet open).

We replaced the time relative to opening indicators with a single indicator for the post period to estimate the difference in difference model:

$$O_{it} = \pi \mathbf{1}[t \geq R_i] + \delta_i + sigma_t + \epsilon_{it}$$

Where π is the coefficient of interest.

In our difference-in-difference models, we assessed changes relative to the reopening date, 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 and our student exposure index. We used Driscoll-Kraay standard errors which account for first-order autocorrelation and clustering of observations at the county level. Results were considered statistically significant if the two-sided p-value was less than 0.05.

B. Limitation of main analysis

As a limitation our data do not contain such measures so we cannot assess the effectiveness of within-college strategies. Similarly, we are unable to test what has occurred once colleges change decisions, such as changing modes temporarily or encouraging students to return home (11). One can also make a strong theoretical argument that sending students home risks further spreading the pandemic across the country by providing new infusions of cases in areas that may relatively low rates of transmission. These concerns have been raised by public health officials, some of whom have publicly opposed closing dormitories, even after a college or university transitioned to online education(12). Further research on the effects of sending students home is needed to understand the risks and benefits of closing residence halls.

Our results are limited by the nature of our data. Our mobility analysis relies on observing cellular GPS signals and these 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 or watch) that provide data under distinct identifiers. Second, we are unable to measure cases among college students vs. others in the county community. Fourth, our measure of mobility does not take account of students that may live in off-campus housing and take classes online.

References

1. Marsicano C. C2I: The College Crisis Initiative – Crisis to Innovation [Internet]. [cited 2020 Sep 14]. Available from: https://collegecrisis.org/

2. Department of Homeland Security. Colleges and Universities Campuses. 2020 [cited 2020 Sep 14]; Available from: https://hifld-geoplatform.opendata.arcgis.com/datasets/colleges-and-universities-campuses?geometry=25.805,-22.656,-25.171,70.117

3. Cook J, Newberger N, Smalling S. The spread of social distancing. Economics Letters [Internet]. 2020 Nov 1 [cited 2020 Sep 20];196:109511. Available from: http://www.sciencedirect.com/science/article/pii/S0165176520303128

4. Weill JA, Stigler M, Deschenes O, Springborn MR. Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. PNAS [Internet]. 2020 Aug 18 [cited 2020 Sep 20];117(33):19658–60. Available from: https://www-pnas-org.libproxy.uncg.edu/content/117/33/19658

5. Holtz D, Zhao M, Benzell SG, Cao CY, Rahimian MA, Yang J, et al. Interdependence and the cost of uncoordinated responses to COVID-19. PNAS [Internet]. 2020 Jul 30 [cited 2020 Jul 31]; Available from: https://www.pnas.org/content/early/2020/07/29/2009522117

6. Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. CSSEGISandData/COVID-19 [Internet]. 2020 [cited 2020 May 29]. Available from: https://github.com/CSSEGISandData/COVID-19

7. Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases - Volume 26, Number 6—June 2020 - Emerging Infectious Diseases journal - CDC. [cited 2020 Sep 14]; Available from: https://wwwnc.cdc.gov/eid/article/26/6/20-0357_article

8. Keeling MJ, Rohani P. Modeling Infectious Diseases in Humans and Animals [Internet]. Princeton University Press; 2011 [cited 2020 Sep 21]. Available from: http://www.jstor.org/stable/10.2307/j.ctvcm4gk0

9. Liu Q-H, Ajelli M, Aleta A, Merler S, Moreno Y, Vespignani A. Measurability of the epidemic reproduction number in data-driven contact networks. Proc Natl Acad Sci USA [Internet]. 2018 Dec 11 [cited 2020 Sep 21];115(50):12680–5. Available from: http://www.pnas.org/lookup/doi/10.1073/pnas.1811115115

10. Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. American Journal of Epidemiology [Internet]. 2013 Nov 1 [cited 2020 Sep 21];178(9):1505–12. Available from: https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwt133

11. Mangrum D, Niekamp P. College student contribution to local COVID-19 spread: Evidence from university spring break timing. Available at SSRN 3606811. 2020;

12. CNN HY Maggie Fox and Jamie Gumbrecht. CDC official affirms coronavirus deaths really are coronavirus deaths [Internet]. CNN. 2020 [cited 2020 Sep 2]. Available from: https://www.cnn.com/2020/09/02/health/us-coronavirus-wednesday/index.html

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

Additional Materials

eTable 1: Association of Reopening and Modality with Mobility and New Cases Using Difference-in-Difference Models

	% change in visitors		Daily new cases per thousand				R_t
	Model 1	Model2	Model 1	Model 2	Model 3	Model 4	Model 4
Open	0.457***		0.017*		0.016*		
	(0.048)		(0.007)		(0.007)		
	[0.362, 0.551]		[0.003, 0.030]		[0.003, 0.030]		
Open, Primarily online		0.328***		-0.001		-0.001	0.019***
		(0.045)		(0.010)		(0.010)	(0.005)
		[0.241, 0.415]		[-0.021, 0.018]		[-0.021, 0.019]	[0.010, 0.028]
Open, Primarily in person		0.539*** c		0.024** a		0.023** a	0.021***
		(0.057)		(0.008)		(0.008)	(0.004)
		[0.427, 0.650]		[0.008, 0.040]		[0.007, 0.038]	[0.013, 0.028]
Open x Student exposure					0.119***		
					(0.021)		
					[0.078, 0.160]		
Open, Primarily online x Student exposure						0.057*	0.011
						(0.028)	(0.017)
						[0.002, 0.111]	[-0.023, 0.045]
Open, Primarily in person x Student exposure						0.142*** a	0.010
						(0.025)	(0.014)
						[0.092, 0.192]	[-0.017, 0.037]

Notes: Sample includes 217,051 Census Block Group for 61 days (% change in visitors) and 3126 counties observed for 61 days for daily new cases per thousand. Student exposure is the change in the average of the log weekly incidence from July 15th to fourteen days before campus reopened, weighted by the change in movement from each county to each campus from the start of the Fall 2019 semester. Standard errors account for first order serial correlation.

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

a Difference from open, primarily online P<0.05

b Difference from open, primarily online P<0.01

c Difference from open, primarily online P<0.001

eTable 2: Association of Reopening with Mobility, Cases, and Disease Transmission by Week

	Visitors	Daily new cases per '000	R_t
1-7 days before opening, primarily online	0.349***	-0.012	0.008*
	(0.037)	(0.008)	(0.004)
	[0.276, 0.422]	[-0.027, 0.003]	[0.001, 0.015]
1-7 days before opening, primarily in-person	0.563*** c	0.001	0.011***
	(0.037)	(0.006)	(0.003)
	[0.491, 0.635]	[-0.010, 0.012]	[0.005, 0.017]
Open for 0-6 days, primarily online	0.412***	-0.000	0.015***
	(0.049)	(0.011)	(0.004)
	[0.317, 0.507]	[-0.021, 0.021]	[0.006, 0.024]
Open for 0-6 days, primarily in-person	0.766*** c	0.008	0.016***
	(0.056)	(0.008)	(0.004)
	[0.655, 0.876]	[-0.006, 0.023]	[0.009, 0.024]
Open for 7+ days, primarily online	0.398***	-0.004	0.024***
	(0.052)	(0.012)	(0.006)
	[0.296, 0.499]	[-0.028, 0.019]	[0.013, 0.035]
Open for 7+ days, primarily in-person	0.633*** c	0.032** a	0.027***
	(0.061)	(0.010)	(0.004)
	[0.513, 0.753]	[0.012, 0.051]	[0.019, 0.036]

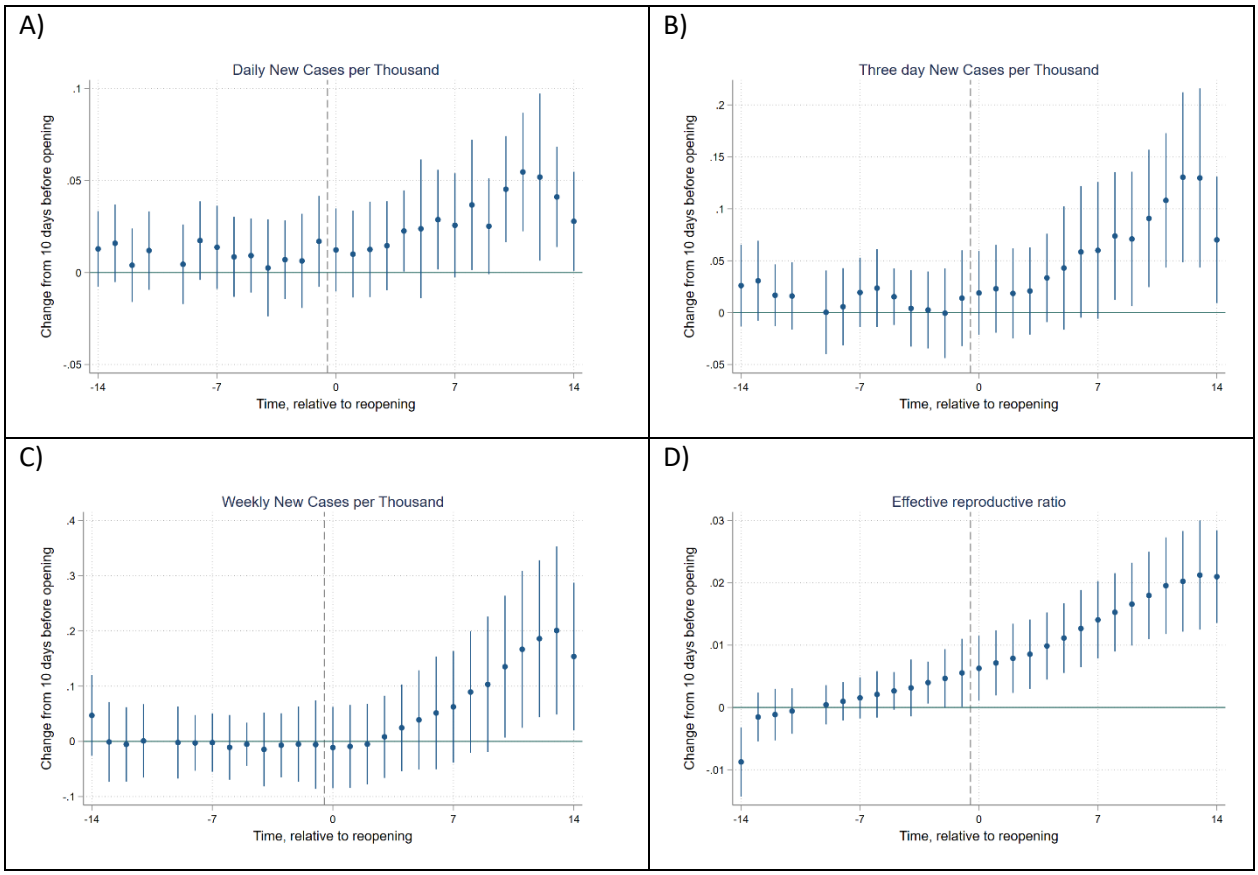
Notes: Sample includes 217,051 Census Block Group for 61 days (% change in visitors) and 3126 counties observed for 61 days for daily new cases per thousand. Standard errors account for first order serial correlation.

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

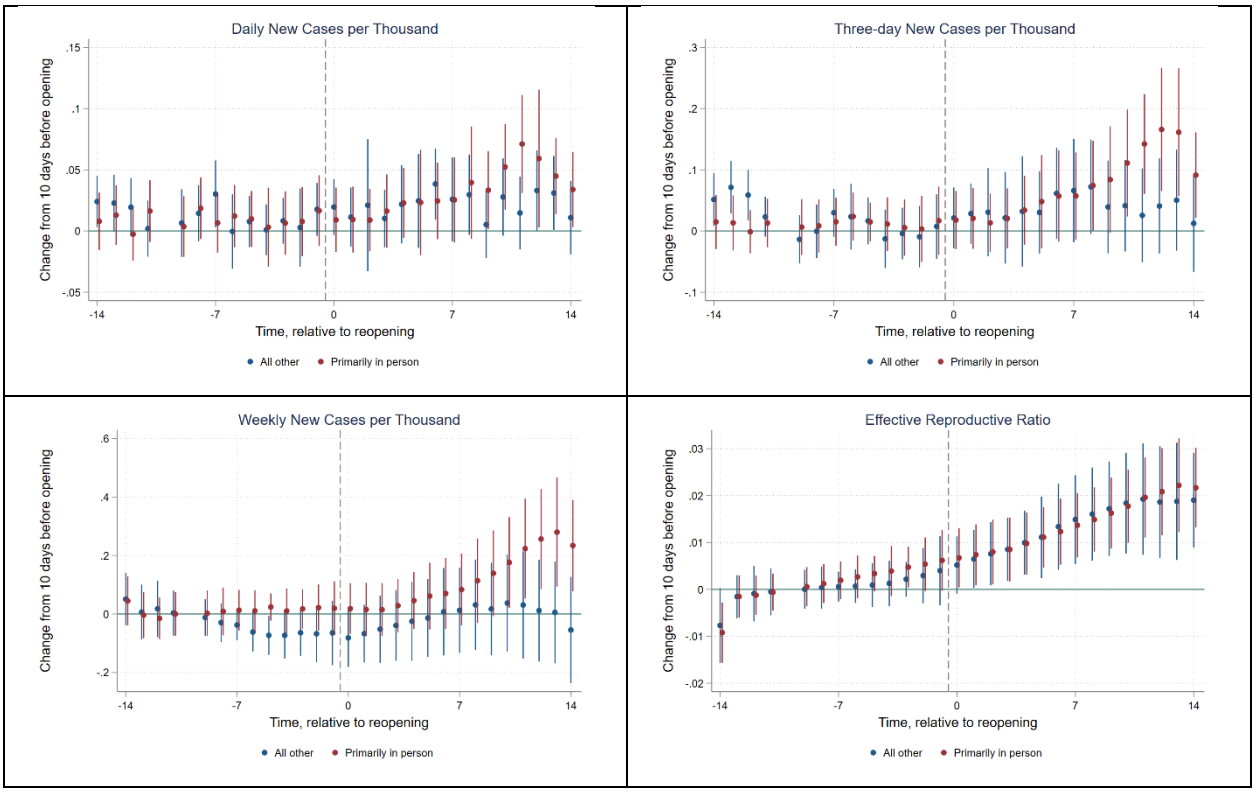
a Difference from open, primarily online P<0.05

b Difference from open, primarily online P<0.01

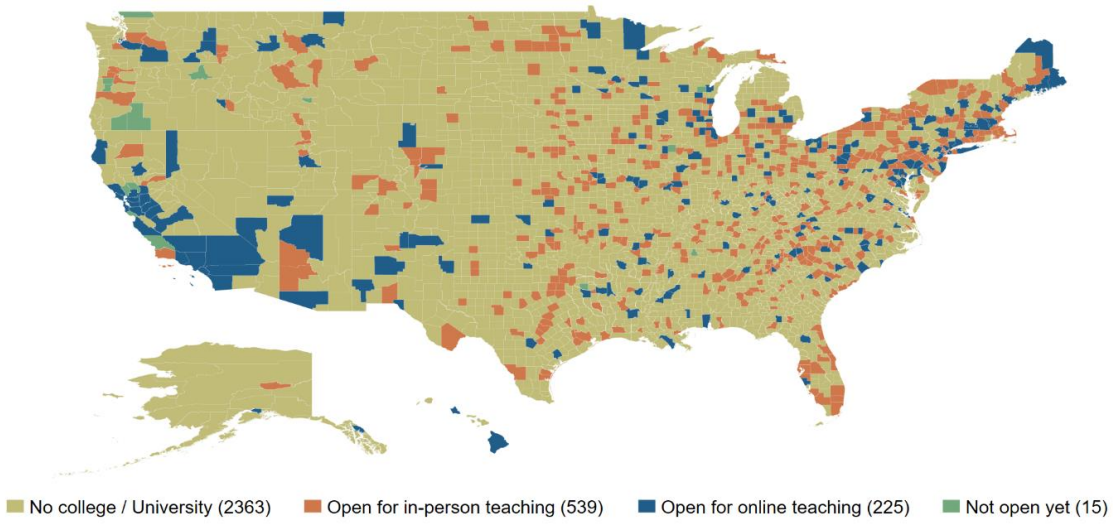
c Difference from open, primarily online P<0.001



eFigure 1: Event studies of trends in daily new cases (A), new cases over the past three days (B), over the past week (C), and R_t (D). New cases are expressed as rates (per thousand county residents). All models include county and date fixed effects, in addition to time relative to reopening. The day ten days before reopening is the omitted reference category.



eFigure 2: Event studies including in-person interactions of trends in daily new cases (A), new cases over the past three days (B), over the past week (C), and R_t (D). New cases are expressed as rates (per thousand county residents). All models include county and date fixed effects, in addition to time relative to reopening. The day ten days before reopening is the omitted reference category.



	School classification		
County classification	Open, primarily in-person	Open, primarily online	Not open yet
No college / University	0	0	0
Open, primarily in-person	761 (83.9%)	131 (14.4%)	15 (1.7%)
Open, primarily online	125 (25.6%)	352 (72.0%)	12 (2.5%)
Not open yet	0	0	13 (100%)
Total	886 (62.9%)	483 (34.3%)	40 (2.8%)

eFigure 3: Mapping schools and counties. Top panel: US map with counties and schools. Bottom panel School classifications

eTable 3: Difference-in-difference estimates of the association between reopening and new cases over the last three days, per 1,000 county residents.

	Model 1	Model 2	Model 3	Model 4
Open	0.044*		0.042*	
	(0.020)		(0.020)	
	[0.004, 0.083]		[0.002, 0.081]	
Open, Primarily Online		-0.009		-0.008
		(0.029)		(0.029)
		[-0.067, 0.048]		[-0.066, 0.049]
Open, Primarily in person		0.065** a		0.061** a
		(0.024)		(0.024)
		[0.019, 0.112]		[0.015, 0.108]
Open x Home state exposure			0.372***	
			(0.062)	
			[0.251, 0.494]	
Open, Primarily Online x Home state exposure				0.188*
				(0.079)
				[0.033, 0.344]
Open, Primarily in person x Home state exposure				0.442*** a
				(0.076)
				[0.293, 0.590]

Notes: Sample includes 3126 counties observed for 61 days. Student exposure is the change in the average of the log weekly incidence from July 15th to fourteen days before campus reopened, weighted by the change in movement from each county to each campus from the start of the Fall 2019 semester. Standard errors account for first order serial correlation.

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

a Difference from open, primarily online P<0.05

b Difference from open, primarily online P<0.01

c Difference from open, primarily online P<0.001

eTable 4: Difference-in-difference estimates of the association between reopening and new cases over the last seven days, per 1,000 county residents.

	Model 1	Model 2	Model 3	Model 4
Open	0.074		0.069	
	(0.047)		(0.047)	
	[-0.017, 0.166]		[-0.023, 0.161]	
Open, Primarily Online		-0.049		-0.046
		(0.065)		(0.065)
		[-0.177, 0.079]		[-0.174, 0.082]
Open, Primarily in person		0.125* a		0.113* a
		(0.055)		(0.055)
		[0.017, 0.233]		[0.005, 0.222]
Open x Home state exposure			0.928***	
			(0.142)	
			[0.649, 1.207]	
Open, Primarily Online x Home state exposure				0.489**
				(0.179)
				[0.139, 0.839]
Open, Primarily in person x Home state exposure				1.093*** a
				(0.175)
				[0.750, 1.437]

Notes: Sample includes 3126 counties observed for 61 days. Student exposure is the change in the average of the log weekly incidence from July 15th to fourteen days before campus reopened, weighted by the change in movement from each county to each campus from the start of the Fall 2019 semester. Standard errors account for first order serial correlation.

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

a Difference from open, primarily online P<0.05

b Difference from open, primarily online P<0.01

c Difference from open, primarily online P<0.001

eTable 5: Effective reproductive ratio (main models): Difference-in-difference estimates of the association between reopening and the effective reproductive ratio (R_t).

	Model 1	Model 2	Model 3	Model 4
Open	0.020***		0.020***	
	(0.003)		(0.003)	
	[0.014, 0.027]		[0.014, 0.027]	
Open, Primarily Online		0.019***		0.019***
		(0.005)		(0.005)
		[0.010, 0.028]		[0.010, 0.028]
Open, Primarily in person		0.021***		0.021***
		(0.004)		(0.004)
		[0.013, 0.028]		[0.013, 0.028]
Open x Home state exposure			0.010	
			(0.011)	
			[-0.012, 0.033]	
Open, Primarily Online x Home state exposure				0.011
				(0.017)
				[-0.023, 0.045]
Open, Primarily in person x Home state exposure				0.010
				(0.014)
				[-0.017, 0.037]

Notes: Sample includes 3126 counties observed for 61 days. Student exposure is the change in the average of the log weekly incidence from July 15th to fourteen days before campus reopened, weighted by the change in movement from each county to each campus from the start of the Fall 2019 semester. Standard errors account for first order serial correlation.

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

a Difference from open, primarily online P<0.05

b Difference from open, primarily online P<0.01

c Difference from open, primarily online P<0.001

eTable 6: Sensitivity of results to weighting counties by population.

	New cases per thousand			Reproductive ratio
	Daily	**Last three days**	**Weekly**	
Open, Primarily Online	-0.017	-0.052	-0.113	0.008
	(0.014)	(0.041)	(0.094)	(0.005)
	[-0.044, 0.009]	[-0.133, 0.029]	[-0.296, 0.071]	[-0.001, 0.018]
Open, Primarily in person	-0.018	-0.057	-0.157	0.005
	(0.014)	(0.043)	(0.097)	(0.004)
	[-0.046, 0.010]	[-0.141, 0.027]	[-0.348, 0.033]	[-0.002, 0.012]
Open, Primarily Online x Home state exposure	0.098**	0.311**	0.737**	0.037*
	(0.034)	(0.105)	(0.239)	(0.016)
	[0.031, 0.166]	[0.105, 0.516]	[0.268, 1.205]	[0.006, 0.068]
Open, Primarily in person x Home state exposure	0.162***	0.495***	1.219***	0.050***
	(0.032)	(0.095)	(0.232)	(0.011)
	[0.100, 0.224]	[0.309, 0.682]	[0.765, 1.674]	[0.028, 0.072]

Notes: Sample includes 3126 counties observed for 61 days. Student exposure is the change in the average of the log weekly incidence from July 15th to fourteen days before campus reopened, weighted by the change in movement from each county to each campus from the start of the Fall 2019 semester. Standard errors account for first order serial correlation.

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

a Difference from open, primarily online P<0.05

b Difference from open, primarily online P<0.01

c Difference from open, primarily online P<0.001

eTable 7: Interactions by C2i Short Mode of Instruction Classification

	New cases per thousand			Reproductive ratio	Visitors
	Daily	**Last three days**	**Weekly**		
Open, Fully In Person	0.045	0.127	0.251	0.029*	0.295***
	(0.027)	(0.080)	(0.187)	(0.013)	(0.054)
	[-0.008, 0.097]	[-0.030, 0.284]	[-0.117, 0.618]	[0.003, 0.054]	[0.190, 0.400]
Open, Primarily In Person	0.026*	0.071*	0.140	0.018***	0.545***
	(0.011)	(0.032)	(0.073)	(0.005)	(0.062)
	[0.005, 0.047]	[0.009, 0.133]	[-0.003, 0.283]	[0.009, 0.027]	[0.424, 0.667]
Open, Hybrid or Hyflex	0.017	0.047	0.079	0.024***	0.555***
	(0.011)	(0.031)	(0.073)	(0.005)	(0.069)
	[-0.003, 0.038]	[-0.015, 0.108]	[-0.063, 0.222]	[0.014, 0.034]	[0.419, 0.690]
Open, Primarily Online	0.001	-0.003	-0.033	0.020***	0.383***
	(0.011)	(0.031)	(0.068)	(0.005)	(0.050)
	[-0.020, 0.022]	[-0.063, 0.058]	[-0.167, 0.101]	[0.011, 0.030]	[0.284, 0.481]
Open, Fully Online	-0.016	-0.058	-0.161	0.011	0.125***
	(0.023)	(0.073)	(0.172)	(0.012)	(0.038)
	[-0.062, 0.030]	[-0.200, 0.085]	[-0.499, 0.176]	[-0.014, 0.035]	[0.051, 0.199]

Notes: Standard errors account for first order serial correlation.

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

a Difference from primarily online P<0.05

b Difference from primarily online P<0.01

c Difference from primarily online P<0.001