Impacts of teaching modality on U.S. COVID-19 spread in fall 2020 semester

Syed Badruddoza

Department of Agricultural and Applied Economics, Texas Tech University

Modhurima Dey Amin

Department of Agricultural and Applied Economics, Texas Tech University

Corresponding author: Syed Badruddoza. TTU Ag Sci, 1312 Boston Ave. Room 204A, Lubbock, TX 79409. Phone: (806) 834-1921. Email: Syed.Badruddoza@ttu.edu

Abstract

We study the impact of college reopening in Fall 2020 on county-level COVID-19 cases and deaths using the information of 1,076 randomly chosen four- and two-year undergraduate degree-granting colleges from the National Center for Education Statistics. These institutions include public, private nonprofit, and for profit schools from 50 US states and the District of Columbia. We match college and county characteristics using several methods and calculate the average treatment effects of three teaching modalities: in-person, online, and hybrid on COVID-19 outcomes up to two months after college reopening. In pairwise comparison, colleges reopened with in-person teaching mode were found to have about 35 percentage points more cases within 15 days of reopening, compared to those that reopened online, and the gap widens over time at a decreasing rate. Death rates follow the pattern with a time lag. However, colleges with hybrid mode reach up to the rates of in-person mode after some time. We also find that greater endowment and student population, bigger class size, and fewer Republican voters in the county are major predictors of choosing remote teaching modes over in-person.

Keywords: College, COVID-19, Online, Teaching modality, United States, University.

JEL codes: A23, I18, I23.

Acknowledgements: Authors would like to thank the Editor and Reviewers, and the participants of the American Economic Association annual meeting 2022 for their helpful comments. Authors have no conflict of interests. No human subject was involved in this research. No funding was received.

1 Introduction

The COVID-19 pandemic and its containment measures have unprecedentedly affected human health and economic activities (e.g., McKibbin and Fernando 2020; Atkeson 2020). Countries across the world have implemented partial or full business and school closures to mitigate the spread of infection (Panovska-Griffiths et al. 2020; Di Domenico et al. 2020). Many U.S. colleges temporarily closed or switched to online in the spring semester, and over six out of ten colleges reopened in the fall semester with an in-person or a combination of in-person and online teaching plans (Gallagher and Palmer 2020; Picault 2021). College students mainly fall in the age cohort of 18 to 29 years, which has a lower death rate (0.4 percent) from COVID-19 (Wrighton and Lawrence 2020; Goodman et al. 2021), but a greater chance of socialization than other age cohorts. Although colleges have taken different measures–including symptom screening, contact tracing, cleaning and disinfecting, and mandatory mask policies–there is a chance that reopening colleges may spread the virus among faculty and staff serving students, as well as in the region around the college (Hubler and Hartocollis 2020).

Given the substantial risk of spread of infection from college campuses to the community, a policy question is, if and to what extent, switching to a more remote teaching modality helps in controlling communicable diseases. Although the severity of the pandemic somewhat reduced by the end of 2021, a study that analyzes the effect of different teaching modalities may offer directions in similar crises in the future—such as an expansion of a new variant of COVID-19. The current study applies a quasi-experimental approach to find the impacts of teaching modality on infection spread onto the neighboring areas. Intuitively, we match college and county characteristics to find a pair of similar colleges, where one chose in-person and the other online. Then we measure how choosing an in-person mode might have affected COVID-19 cases and deaths in the neighborhood as

opposed to choosing an online mode. We perform similar exercises for two other pairs: in-person versus hybrid, and hybrid versus online. Results provide evidence that college reopening, especially within in-person teaching mode, increases COVID-19 cases and deaths over time.

Most studies so far have looked into the impact of elementary and secondary school closure, optimal reopening plans, and teaching strategy (e.g., Di Domenico et al. 2020; Panovska-Griffiths et al. 2020; Gandolfi 2020; Cohen et al. 2020). Walke et al. (2020) discuss the possible issues of U.S. college reopening and the prevention measures required on campuses. Some early studies find an aggregated negative effect of school closures on COVID cases and deaths in the community (e.g., Auger et al. 2020; Pan et al. 2020), whereas some others find little to no negative effect (e.g., UNICEF 2020; Herby et al. 2022). A more granular level study conducted by Chernozhukov et al. (2021) shows that in-person opening of K-12 schools relative to that with remote opening is associated with a 5 percentage point increase in the weekly growth rate of cases. However, remote teaching in colleges may relate to COVID-19 containment more than remote teaching in schools, because college students fall into age groups that are more vulnerable to the disease than their younger counterparts (Goodman et al. 2021). In a study close to ours, Andersen et al. (2020) use an event study under a difference-in-difference setting to investigate college openings' association with cellphone-tracked human mobility and COVID-19 cases in the counties of the campuses. Their preliminary results suggest that, reopening increased county COVID-19 cases by 1.7 daily per 100,000 county residents, and in-person teaching created greater visits and higher cases during the reopening weeks, but no contribution from online teaching was found. Nevertheless, most of these studies do not consider the fact that the choice of reopening and teaching modality is endogenous to observed cases. One of the assumptions of the difference-in-difference method is that the allocation of intervention was not determined by outcome, which is violated in this case because teaching modalities were often chosen by colleges based on pre-existing COVID rates in the region.

Moreover, COVID cases can be influenced by required testings during the reopening weeks, and deaths may take longer than two weeks to appear in the data. Additionally, there can be other institutions (e.g., K-12 schools) affecting both COVID outcomes and the decision of instructional modalities. Our primary source of identification is that different colleges reopened on different dates, allowing for a clear comparison. We utilize data from various sources, including cellphone mobility data for comprehensive matching; and we use methods that account for multidimensional differences among colleges and re-weight the sample based on the college's possibility of selecting an instructional modality.

2 Method

We are interested in finding the effects of teaching modality on two COVID-19 outcomes: (1) new COVID-19 cases reported in the county, and (2) new COVID-related deaths in the county. The impact of teaching modality is assessed on county-level outcomes (and not on the college level) because students may spread the virus to non-students and immunocompromised people in the community, which is a critical policy concern. First, we specify a simple relationship between college reopening in the fall semester and the outcome as below,

$$Y_i = \text{Intercept} + \alpha \text{T}_i + \text{Controls}_i \times \beta + \text{error}_i \#(1)$$

where, Y_i is a COVID-19 outcome, $i = 1, \ldots, N$ indicate colleges and T_i is a binary treatment variable that represents teaching modality. We use three specifications of T_i separately: (1) in-person=1, online=0, (2) in-person=1, hybrid=0, and (3) hybrid=1, online=0 for pairwise

comparison of teaching modalities. That is, for each of three regressions we only include the observations with respective teaching modalities. The vector of control variables includes observed college- and county-level covariates such as college endowments or percent of the population who stayed at home, and β is a vector of respective parameters. Our parameter of interest is α, which shows the impact of a mode of instruction on an outcome variable, e.g., whether switching from online to in-person has an impact on the COVID-19 outcomes. The equation above would have been identified if colleges were homogeneous and were randomly assigned a teaching mode (Barnow et al. 1981). In practice, the treatment assignment is not random, because colleges choose the mode of teaching based on their distance education capacities, observed COVID-19 cases, and many other characteristics. The probability of adopting a teaching modality may vary considerably across colleges. Intuitively, if we could select two colleges with identical institutional and regional features, except one went in-person and the other online, we could measure the difference in COVID outcomes induced by the teaching modality.

There are some other estimation issues. First, many areas had high COVID-19 cases in the spring and summer semesters. Due to the exponential nature of the disease spread, places with more initial cases will experience faster growth. Moreover, teaching modality in the fall semester may depend on existing cases, thus generating a reverse causality. Second, there can be observed and unobserved college and county features that affect both teaching modality and COVID-19 outcomes, creating an omitted variable bias. Third, a measurement error may occur because infections are reported less during the weekend and more on weekdays, and the incubation or survival period varies. In addition, colleges reopening with in-person teaching elements may require mandatory COVID testing, which leads to an increase in the number of cases. Finally, a county may have multiple colleges or K-12 schools that determine the presence of people in the region, which in turn may affect both college teaching modality and COVID spread. We take several measures to

address these problems. For instance, we control for cellphone mobility, many observed college and county characteristics, and use percentage changes of the outcome variables to eliminate college and county-level unobserved temporal fixed effects. We also check the outcomes for up to two months and control for the presence of other colleges in a county. Attempts to address other estimation issues are described below[1].

One source of identification is that different colleges reopened on different dates between July and October of 2020. Therefore, we can isolate the effect of reopening on COVID-19 from the aggregate trend. A bigger identification problem is the presence of other colleges, K-12 schools, clubs, churches, etc. in the county that are unobserved in the data. Trips to these areas may affect the disease spread, as well as the selection of college teaching modalities. Thus, omitting the level of crowdedness or trips made in the county will lead to omitted variable bias. We follow Chernozhukov et al. (2021) to address the problem. They used foot traffic in K-12 schools captured by cellphone mobility data to measure if the school is operating in-person (more foot traffic) or online (less foot traffic). Notice that, even if we control for the foot traffic in K-12 schools, we still need to control for the foot traffic in numerous other places to avoid the omitted variable bias, i.e., we must add many other variables in the regression model. A parsimonious way to address the issue is to control for the percentage of the county population who stayed home, recorded by their cellphone locations, according to the following logic:

$$\text{Total population} = \text{population at home} + \text{population outside home} \quad \# (2)$$

[1] The motivation of addressing the estimation issues is to try our best to make sure that the treatment assignment was, at least statistically, "equally likely" and the estimates are not biased. A randomized placebo-controlled trail would have generated a more accurate estimate. We discuss this limitation in the concluding section.

$$\text{Population outside home} = \text{foot traffic in colleges} + \text{foot traffic in K12 schools} + \ldots \# \ (3)$$

$$\text{Population at home (\%)} = 100 \times \frac{\text{total population} - \text{population outside home}}{\text{total population}} \# \ (4)$$

where, being outside is specified by staying outside the home for longer than ten minutes[2]. Thus, controlling for the percentage of the population who stayed home benefits the estimation process in many ways, because it comprehensively covers people who are running errands, going for a COVID test, going to schools, colleges, and any other places. We give two examples below.

First, if there are multiple colleges in the county and we do not observe whether the unobserved colleges are in-person or online, cellphone mobility captures the variation. If the unobserved college reopens in-person, average cellphone mobility in the county should increase more than if it reopens online. On the flip side, the percentage of the population staying home will decrease (eq. 4). Then the increase in COVID outcomes will be explained by the percentage of the population staying outside in various places, and will not corrupt the coefficient with the teaching modality for the observed college. Moreover, if the college teaches in-person but follows a strict "closed" campus policy, cellphone mobility captures that as well. Assume two colleges both teaching online, but one has more students returned to campus for an open campus policy. The cellphone mobility controls for the variation in COVID outcomes unexplained by teaching modality.

[2] Cellphone locations are accurate up to 30 meters on average, and are recorded in every 10 minutes for 24/7. The percentage of people went outside home means people who went farther than 30 meters from their night-staying location at any point in 24 hours for longer than 10 minutes. The implicit assumption we make here is at least college students have access to a cellphone device. This is not unrealistic because Pew Research Center (2021) reports that 100% of Americans aged 18-29 own a cellphone of some kind. Another assumption is to define fixed night-staying location as home.

Second, Chernozukov et al. (2021) found a 0.09 correlation between K-12 school visits and college visits during the reopening weeks. Thus, if K-12 school reopening is correlated with college teaching modality, our estimation will be biased. Following Chernozukov et al. (2021), the reopening of K-12 schools will reflect in their foot traffic, and controlling for cellphone mobility will capture the unobserved variation in COVID outcomes, which allows us to estimate the effect of college teaching modality. Infection rates in neighboring counties will affect a county via population mobility, so a splillover effect will be controlled for as well.

What does the coefficient with the teaching modality variable capture? Assume two comparable colleges in two counties have the same foot traffic, and counties are similar in all other aspects related to COVID and college teaching modalities, both of the colleges have students on campus and similar average cellphone mobility per day, but one of them reopened online and the other in-person. Our estimate gives the difference between covid cases in two counties generated by the teaching modalities only: the effect of students attending an in-person class as opposed to an online one, holding everything else fixed.

To deal with the heterogeneous starting values of COVID, measurement errors (e.g., more testing on weekdays), and threshold selection issues (e.g., incubation period), we use two-week intervals: 0-15 days after reopening, 15-30 days after, 30-45 days after, and 45-60 days after reopening; and run separate cross-section regressions for each subsample. The two-week period is selected based on the incubation periods (e.g., Paul and Lorin 2021). The benchmark of percentage calculation is 0-15 days before reopening; for example,

$$\%\ \text{change in cases}_{\text{0 to 15 days after reopening}} =$$

$$100 \times \frac{\text{cases}_{\text{0 to 15 days after reopening}} - \text{cases}_{\text{0 to 15 days before reopening}}}{\text{cases}_{\text{0 to 15 days before reopening}}} \quad \# \ (5)$$

and so on. We chose two weeks intervals, because doing so gives COVID-19 enough time to appear in the outcomes, and we did not go beyond 60 days because the effects may be contaminated by other exogenous factors.

More testing will result in more cases. The 15-days interval approach, combined with the cellphone mobility above, helps us isolate the effect of increased COVID testing during reopening weeks in colleges with in-person modes. This is intuitive because, (1) testing requirements for college reopening are less likely to go beyond the first month after the semester starts, so a statistically significant α beyond the first 30 days is more likely to represent the contribution of teaching modality, (2) cellphone mobility reflects trips to the testing center as well, hence helps to identify α, and (3) COVID-related deaths are not affected by testing requirements, and incubation-to-death period may take longer than a month. So a positive significant α for COVID-related deaths after 30 days presents a more convincing case.

The next challenge is to ensure the comparability between colleges. We implement various matching methods described below to address this problem. We employ five methods and check if the results are robust: (1) bivariate Ordinary Least Squares (OLS) where $\beta = 0$ in the model above, (2) multivariate OLS where $\beta \neq 0$, (3) propensity score matching (PSM), (4) Nearest Neighbor matching (NN), (5) and Kernel Multivariate Distance Matching (KMDM). The first two methods are conventional. The latter three methods focus on finding "statistical twins" for each observation in the treatment group from the control group with similar values of the covariates. The average treatment effects can be calculated as the mean of differences between the observed values in the

treatment group and the imputed counterfactual values. We present a non-technical version of the matching methods below.

Rosenbaum and Rubin (1983, 1984) discuss propensity score matching (PSM) to adjust the probabilities for the differences in pre-treatment variables. A propensity score is the conditional probability of receiving the treatment given the pretreatment variables, i.e., $P(T_i = 1|X_i)$. The probabilities of treatments generated are used to create weights that adjust the pre-treatment imbalances. For example, if a college has a high probability of choosing online than in-person, greater weights are assigned to in-person than online during the estimation. PSM is usually calculated using logistic regression. For a recent technical discussion, see Imbens (2000) and McCaffrey et al. (2013).

Although finding a match with PSM is relatively easier, many studies argue that matching with propensity scores can be misleading as it ignores the multidimensional differences between two observations and simplifies them into one dimension—the score (e.g., King and Nielsen 2019). The nearest neighbor (NN) matching adopts a more multidimensional approach and uses some (default being one) closest observations in the control group. In the case of ties, NN uses all ties as matches. Conversely, a single control observation can be used many times with replacement. Matching with more than one continuous covariate may induce bias, so we apply a bias-adjusted NN matching proposed by Abadie and Imbens (2011).

Our final matching technique, KMDM, is similar to NN, but uses a certain bandwidth for multivariate matching instead of comparing with the closest neighbor. Simulation studies show that KMDM tends to outperform PSM because it approximates fully blocked randomization that is relatively more efficient (King and Nielsen 2019). We apply KMDM where observations in treatment and control groups are matched using Mahalanobis distance, and use the Epanechnikov

kernel function to assign larger weights to controls with smaller distances. Following Huber et al. (2015), we choose 1.5 times the 90 percent quantile of the non-zero distances in finding a pair. The use of KMDM is more appropriate in our case, given the multidimensional heterogeneity among colleges and counties. See Jann (2017) for further discussion on the advantages of KMDM over other methods.

3 Data and descriptive statistics

For college characteristics, we utilize the Integrated Postsecondary Education Data System (IPEDS) survey annually released by the National Center for Education Statistics (NCES 2020). The IPEDS survey includes information on four- and two-year undergraduate degree-granting colleges and universities from all 50 U.S. states and the District of Columbia. Variables include tuition and fees, enrollment, number of degrees and certificates conferred, number of employees, financial statistics, and so on. Over six thousand such Title IV institutions (i.e., institutions that process U.S. federal student aid) were listed by the NCES in 2020. These institutions include public, private nonprofit, and private for profit schools that provide postsecondary education or training beyond the high school level. We aimed at selecting a thousand colleges for our analysis and randomly chose 1,100 of them using a random number generator under uniform distribution. The extra 100 colleges were selected to ensure that dropping missing values did not decrease the sample size below 1,000. Merging this dataset with other datasets, as described below, eventually gives 1,076 colleges after dropping the missing observations. Among the selected colleges, 41 percent are public, 42 percent are private non-profit, and the remaining 17 percent are private for profit.

Data on teaching modality and the official start date of the fall 2020 semester were manually collected between July and December of 2020 by authors from respective college websites.[3] Over 90 percent of the colleges mentioned their latest teaching modality between June and September 2020. Information for the remaining colleges was obtained from contemporary campus news or local news. Some examples of search phrases we used are, "XXX university fall 2020 covid information reopening plan new" or "XXX college restarts in-person class in fall 2020" or "XXX institute fall 2020 plan president announcement updated". We categorize fall reopening plans into three types: (1) in-person, (2) online, and (3) hybrid. Colleges in the "in-person" group started the fall 2020 semester with face-to-face classes and open residence halls, and may include few online delivery materials; colleges in the "online" group primarily offered online classes with some exceptions for lab components and may have some students on campus; and colleges in the "hybrid" group either divided the class into online and in-person section, or switched the teaching mode on a rolling basis, or offered courses with both in-person and online access. Thus, colleges that changed teaching modalities during the sampled period were considered hybrids. We understand this might create noise in the hybrid category, but we kept the in-person and online categories pure (either full in-person or fully online), as their difference is the main interest of this research. Moreover, the definition of hybrid substantially varies from one college to another, so it is more convenient to utilize in-person and online modes to infer the effects of hybrid modes than to study all types of hybrid modes. In our sample, 406 colleges (37.73 percent) taught online, 386 (35.87 percent) taught in person, and the remaining 284 (26.39 percent) followed a hybrid method.

We obtain daily new COVID-19 cases and deaths at the county level from the New York Times (NYTimes 2020) to stay consistent and comparable with similar studies conducted before

[3] A list of colleges and their webpages are publicly available in school_info.txt on https://github.com.

(e.g., Fox et al. 2020; Andersen et al. 2020; Chernozhukov et al. 2021). Moreover, the NY Times data and CDC data are highly correlated (Chernozhukov et al. 2021). Figure 1 simply plots the changes in COVID-19 cases and deaths as a percent of 15 pre-semester days grouped by college teaching modality. We took the percentage changes to remove the influence of initial values, and to get rid of the college and county-level idiosyncratic features. Different colleges started the fall 2020 semester on different dates from the end of July to October. The earliest reopening date for the fall semester is July 20th in our sample whereas the latest date is October 24th. Figure 1 matches the reopening dates, and suggests that COVID-19 cases were exponentially growing in the fall semester regardless of teaching modality, and COVID-related deaths followed the pattern. However, colleges teaching online are located in counties that had slower growth of cases after the college reopening date, on average. Cases were slowly growing for colleges with hybrid teaching modes as opposed to colleges with in-person modes. After 30 days, the pattern of cases is reflected in the pattern of deaths. On average, counties where colleges taught in hybrid or in-person mode had greater COVID-related deaths than counties where colleges taught online. However, an econometric analysis is required to find the impact of teaching modality on COVID outcomes.

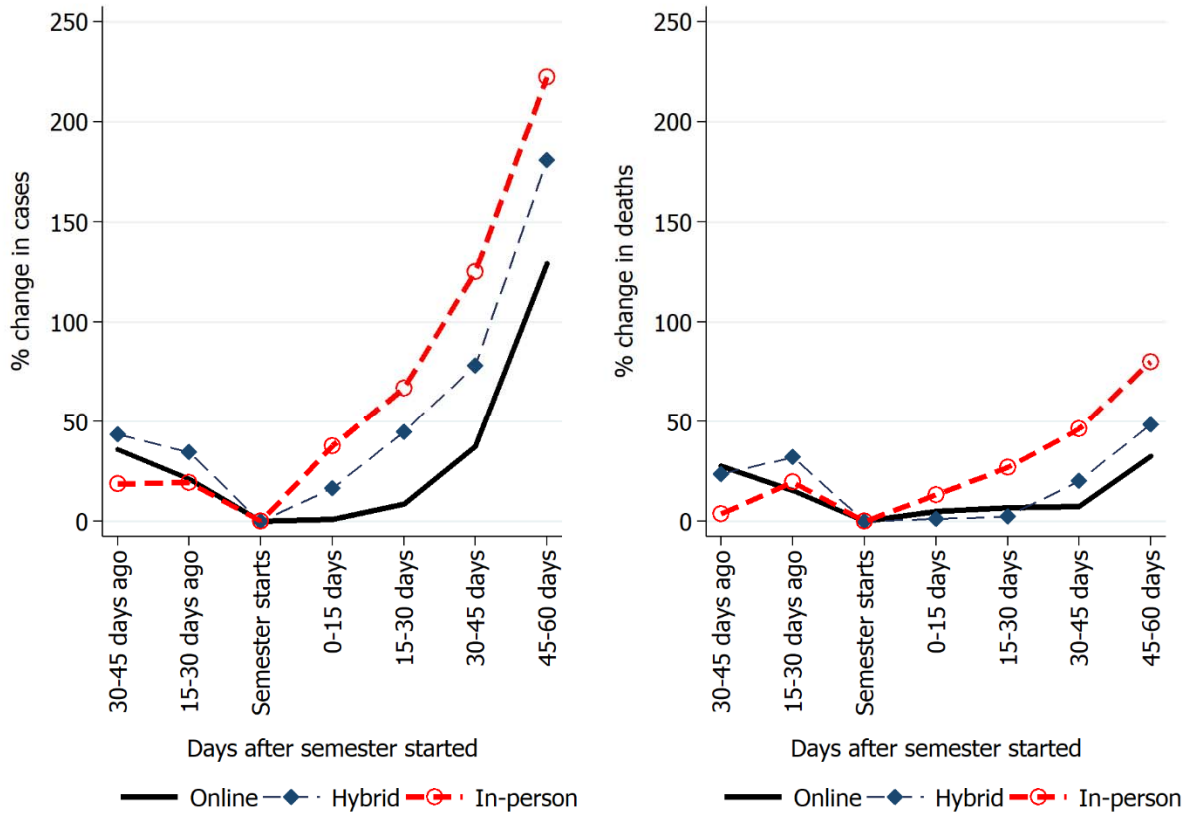

Figure 1. Growth of COVID-19 cases by college teaching modality in fall 2020 semester

Note: Figures are averages across colleges. Daily data of 1,076 colleges from July to October 2020. See the data section for details.

Source: COVID outcomes are obtained from the New York Times (2020), and teaching modalities were collected by authors from the respective college websites.

We select several control variables for the econometric analysis based on the literature. Table 1 provides variable descriptions and sources, and Table 2 presents summary statistics of variables. As discussed in the method section, one of the most important control variables is the average percentage of the county population that stayed home. We collect this variable from the Bureau of Transportation Statistics (BTS 2020). The variable records the seven-day moving average of cellphone mobility at the county level, such that "staying home" is defined by cellphones that do not

go outside the home for longer than 10 minutes. Intuitively, we get a percentage of county population that are not home for each ten minutes, then average the figures to get one number for a day, then calculate the seven-day moving averages. As shown in Table 2, fewer percentages of the county population (24.21 percent) stayed home under the in-person category, compared to hybrid (25.05 percent) and online (26.41 percent). The similar staying home rates across teaching modalities makes the covariate balancing more appropriate. Omitting the stay-home variable makes the effects of a teaching mode on COVID-19 outcomes biased, because COVID-19 outcomes may increase if fewer people stay home, given that cellphone mobility is associated with teaching modalities in schools (Chernozhukov et al. 2021).

We include five college-level controls: (1) enrollment, (2) cost of attendance, (3) endowment per student, (4) student-faculty ratio, and (5) public versus private dummy. The total enrollment variable controls for the size of the college, which can be an important determinant of disease spread (e.g., Fox et al. 2020). Endowment per student and cost of attendance respectively capture the financial capacity of the college and the average affluence of its students, both are critical determinants of technology adoption in schools (e.g., Brahmasrene and Lee 2012; Leu et al. 2015; Sun and Chen 2016; Gallagher and Palmer, 2020). The variable on student-faculty ratio is an indirect measure of the average class size (De Paola et al. 2013), that may affect both case spread and remote teaching decisions. We also include a binary variable on whether the college is public or not because public colleges can be subject to COVID containment rules and regulations. Moreover, many public schools may have administrative or financial constraints to abruptly change teaching modes. Table 2 indicates that these college characteristics vary across teaching modalities.

Control variables on county characteristics were extracted from the American Community Survey (ACS 2020). One of the most important variables is the total population, which primarily

determines the rate of disease spread (Siedner et al. 2020). We further added the percentage of Black population and the percentage of men in the county to capture their disproportionately higher COVID rates (e.g., Millett et al. 2020; Kim et al. 2021). Real household income controls for the average financial capacity of the county population, including students, to adopt new technologies and deal with income shocks during COVID (e.g., Tan et al. 2021). The binary variable on rural campus is added following the definition of US census because the rural location may affect the capacity and decisions regarding teaching modality. A variable on the existing mask ordinance was also included because mandatory mask mandates were found to have a negative association with case growth (Chernozhukov et al. 2021). Individual mask policies in county and state were obtained from HealthData.gov (2020). For some states, decisions regarding public university reopening were determined by central agencies, so we include state fixed effects in all models.

We also added information on county-level shares of Republican votes in Presidential Election 2016 from the MIT Election Data and Science Lab (MIT 2018) to take residents' perception of COVID-19 risk into account (e.g., Tyson 2020). Two other critical county-specific variables are COVID-19 cases and deaths in the spring 2020 semester, because it was the last regular semester that gave the college administration a signal about the disease situation in the community. Higher cases and deaths in spring while the campus was open may influence college administrations to switch to online teaching modes (Gallagher and Palmer 2020). Some studies find an association between temperature and COVID-19 spread (e.g., Livadiotis 2020), so the average temperature in the county during the reopening month was included in the model from the National Oceanic and Atmospheric Administration (NOAA 2021) data. Finally, the total number of colleges in the respective county (including the sampled college) was included in the regression from NCES data, in order to remove the spillover effects, i.e., the contribution of other colleges to COVID outcomes. We admit that the data may not include all offices and institutions that may affect both COVID

cases and teaching modality, but controlling for cellphone mobility together with college count and other identification strategies above should isolate the influence of relevant institutes from our analysis.

A required condition for estimating treatment effects is distributional similarities of pre-treatment covariates across the treatment groups (McCaffrey et al. 2013). Table 2 indicates considerable variation across teaching groups. The difference justifies the use of matching in our analysis. Tests for the distributional equality of matching variables between the treatment and control groups are placed in the appendix (see Figure A1). For PSM, these control variables are used to find the determinants of teaching modality using logistic regressions. The matching process should only contain variables that are measured at baseline, because variables measured at around the treatment may be influenced or modified by treatment (Austin 2011). Therefore, we selected the determinants from the latest academic year available before the fall 2020 semester.

Table 1. Variable description

(1) Variable	(2) Unit	(3) Source	(4) Details
Online	Binary	College websites	Takes 1 if the college holds classes primarily online for fall 2020, zero otherwise
Hybrid	Binary	College websites	Takes 1 if the college adopts a combination of online and in-person modes for fall 2020, zero otherwise
In-person	Binary	College websites	Takes 1 if the college holds classes primarily in person for fall 2020, zero otherwise
New cases	Numeric	NYTimes (2020)	Daily new COVID-19 cases per 100,000 population in the county
New deaths	Numeric	NYTimes (2020)	Daily new COVID-19 deaths per 100,000 population in the county
% stayed home	Percentage	BTS (2020)	Percentage of county population that did not go outside home for longer than 10 minutes
Enrollment	Count	NCES (2020)	Total student count in fall 2018 semester, including graduate and remote students
Cost of attendance	USD	NCES (2020)	Total price for in-state students living on campus for 2019-20
Endowment per student	USD	NCES (2020)	Endowment assets per FTE enrollment at the end of 2018
Student-faculty ratio	Ratio	NCES (2020)	Enrollment divided by total full-time instructional staff
Public	Binary	NCES (2020)	Takes 1 if the institution is primarily funded by a state government, zero otherwise
Total population	Count	ACS (2020)	Total population in the county where the college is located (2018)
Black population	Percentage	ACS (2020)	Percentage of African American population in the county (2018)
Male population	Percentage	ACS (2020)	Percentage of male population in the county (2018)
Rural campus	Binary	College websites	Takes 1 if the college has a rural campus by 2010 census definition, zero otherwise
Household income	USD	ACS (2020)	Median real household income in 2018
Republican votes	Percentage	MIT (2018)	Percentage of votes for the Republican Party at the county level in Presidential election 2016
Mask ordinance	Binary	HealthData.gov (2020)	The county has an ongoing mandatory mask policy for individuals in public places
Cases in spring 2020	Count	NYTimes (2020)	Total COVID-19 cases in the county by May 15, 2020
Deaths in spring 2020	Count	NYTimes (2020)	Total COVID-19 deaths in the county by May 15, 2020
Temperature	Fahrenheit	NOAA (2021)	County mean temperature in reopening month
County college count	Count	NCES (2020)	Number of colleges listed in the NCES data

Note: Means and standard deviations of the variables are placed in Table 2.

Table 2. Summary statistics

(1) Variable	(2) Online	(3) Hybrid	(4) In-person	(5) Overall
Daily new cases	247.00	100.04	72.43	145.59
	(558.23)	(282.69)	(248.92)	(408.97)
Daily new deaths	4.12	1.62	1.09	2.38
	(10.54)	(5.98)	(4.39)	(7.75)
Online				0.38
				(0.48)
Hybrid				0.26
				(0.44)
In-person				0.36
				(0.48)
% stayed home	26.41	25.05	24.21	25.27
	(4.06)	(4.33)	(4.00)	(4.22)
Enrollment	8,828.61	6,307.21	4,417.95	6,580.85
	(9,991.36)	(8,272.99)	(6,972.39)	(8,754.75)
Cost of attendance	44,980.83	46,178.28	41,089.31	43,757.31
	(21,312.56)	(19,431.71)	(16,628.42)	(19,059.39)
Endowment per student	61,279.10	28,545.41	59,892.61	51,811.22
	(347,471.23)	(62,887.43)	(367,714.37)	(307,016.55)
Student-faculty ratio	41.71	28.11	29.18	33.61
	(30.23)	(22.71)	(25.53)	(27.44)
Public	0.60	0.38	0.33	0.44
	(0.49)	(0.49)	(0.47)	(0.50)
Total population	1,578,217	616,050	444,525	917,566
	(2,620,932)	(973,588)	(1,028,931)	(1,867,046)
Black population	14.57	13.68	11.89	13.37
	(15.77)	(14.44)	(13.14)	(14.55)
Male population	49.13	49.13	49.14	49.14
	(1.28)	(1.19)	(1.27)	(1.26)
Household income	64,587.54	59,805.99	56,796.75	60,526.88
	(17,710.66)	(15,156.93)	(14,395.72)	(16,253.30)
Republican votes	38.94	46.09	52.97	45.86
	(16.72)	(16.25)	(14.53)	(16.93)
Mask ordinance	0.86	0.80	0.71	0.79
	(0.35)	(0.40)	(0.45)	(0.41)
Rural campus	0.16	0.22	0.35	0.24
	(0.37)	(0.41)	(0.48)	(0.43)
Cases in spring 2020	12,706.16	5,356.90	3,576.15	7,498.42
	(23,007.05)	(10,396.52)	(10,177.45)	(16,800.38)
Deaths in spring 2020	594.02	288.29	70.89	370.25
	(1,038.35)	(568.51)	(5.298)	(789.65)
Temperature (F) in reopening month	72.37	71.38	70.89	71.58
	(6.14)	(6.07)	(5.30)	(5.86)

County college count	6.697	3.637	3.078	4.591
	(8.462)	(3.579)	(3.748)	(6.174)
Observations (Colleges)	406	284	386	1,076

Note: Figures show averages across colleges. Standard deviations are placed in parentheses. Table 1 provides variable descriptions.

4 Results

Tables 3 and 4 respectively show the average treatment effects of teaching modalities on COVID-19 cases and deaths. We begin with simple bivariate OLS to test the correlation, then control for other covariates, and then use three matching models. Apart from the bivariate OLS (Model 1), all control variables were used in all estimates (Models 2-5).

Table 3 suggests that the association of in-person teaching with COVID-19 cases, as opposed to online, is positive and significant. For example, the bivariate model shows that cases increase by 18.29 percentage points in the first 15 days after reopening compared to the 15 pre-semester days. However, all other treatments and later periods are not statistically significant in the bivariate model. The gap between the effects of in-person and online stays significant when we control for other covariates, using multiple OLS, PSM, NN, and KMDM respectively; but becomes insignificant for multiple OLS and PSM as we move beyond 30 days. A significant effect in later periods is important from a policy perspective because the difference in cases in the first 15 days can be influenced by mandated testing for in-person teaching mode. Matching with propensity scores does not offer a significant advantage over multiple OLS regression. Matching with NN somewhat corroborates the result and indicates a 48 percentage points increase in two months after reopening. The gap between in-person and hybrid is similar (48 percentage points) by the first two months of reopening, as indicated by the NN model (Table 3).

However, PSM collapses all data dimensions into one, whereas NN is vulnerable to the availability of closest neighbors on all dimensions. A more appropriate and robust method for our analysis is KMDM because it separately performs matching on each dimension. Estimates from NN in Table 3 suggest a statistical difference in COVID-19 cases between in-person and online groups for all periods. For instance, a college reopened with an in-person teaching mode may have 54 percentage points more cases in the county than the pre-opening 15 days in the first month, compared to a similar college with an online teaching mode. The figure becomes 81 percentage points after the first two months.

Interestingly, KMDM results show that the hybrid teaching mode has a positive significant difference from online in later periods. Although the gap between hybrid and online is insignificant in the first month, colleges with hybrid modes have 53 percentage points increase in cases as opposed to colleges with online modes by the end of the second month. The gap between in-person and hybrid remains statistically insignificant. That is, the hybrid mode follows the pattern of in-person after 30 days. To summarize, colleges that chose in-person are associated with a greater increase in COVID-19 cases in their respective counties, compared to the colleges that chose online. A hybrid instructional mode may create a difference from in-person at the beginning, but may not sustain for longer periods, possibly because the spread from in-person components overwhelms the gain from online components due to the exponential nature of the disease spread.

Table 4 presents estimation results for COVID-19 deaths. Relatively simpler models (1-3) show some positive significant differences between in-person and hybrid within the first month. It is possible that colleges with high pre-semester cases influenced the death rates within 30 days when control variables were not properly matched. However, as a more robust method, KMDM captures a statistically significant difference between in-person and online modes, which grows over time at a

decreasing rate. For example, teaching in-person as opposed to online is associated with 29 percentage points more deaths in the first 30 days after reopening compared to the 15 pre-semester days, and 38 percentage points more deaths in the first 60 days after reopening. Similar to the results we found for COVID cases, hybrid modes create a statistically significant difference in death rates from in-person for the first 45 days, but the difference gradually dissolves beyond that period.

In short, the magnitude of the average treatment effects increases at a decreasing rate as we go from 0-15 days to 45-60 days in both Tables 3 and 4. This occurs because college reopening increases average cases and deaths regardless of teaching modality. Although online mode raises the rates slower than in-person, the gap starts to shrink after 45 days, possibly because college students are likely to socialize on campus outside the classroom even if the classes are online.

The dataset further allows us to explore another aspect of teaching modality during a pandemic. We use logistic regressions to find the predictors of a teaching modality. The results in Table 5 show that colleges with greater enrollment are less likely to choose in-person than online (column 1), and in-person than hybrid (column 3), holding other variables constant. A percent increase in endowment per student decreases the log odds ratio of choosing in-person over online by 0.616, and that of choosing hybrid over online by 0.476, given the other predictors. More population in the county might have made colleges choose remote teaching components. For example, colleges in counties with a percent more population are associated with 0.812 lower log odds of choosing in-person modes over online. The negative association of in-person mode with enrollment and population variable implies that the colleges might have considered remote teaching elements where campuses are more likely to be crowded.

Colleges with bigger class sizes, measured by the student-faculty ratio, were less likely to choose in-person teaching modes. The log odds ratio of choosing an in-person mode over online

decreases by 0.82 as the student-faculty ratio increases by one percent, whereas that of hybrid decreases by 1.125 over online for the same increase in student-faculty ratio. Table 5 also shows that greater Republican votes in the 2016 election have a positive association with choosing in-person mode over online or hybrid. As discussed in the data section, political views in a county might have played a role in the selection of teaching modality. Interestingly, the real household income in the county has a positive relationship with in-person teaching mode as compared to online. Since college-level financial variables are already controlled for, a possible explanation for this is that wealthier counties preferred in-person over online to reduce the stress and to ensure the quality of education (e.g., Lazarevic and Bentz 2021; Kofoed et al. 2021; Orlov et al. 2021).

The above results are based on the full model that was constructed based on the literature. Some control variables can be correlated, which reduces the precision of estimated coefficients in the logistic model. Figure A2 in the appendix presents the correlation heatmap of control variables. Not surprisingly, total population, COVID-19 cases and deaths in the spring semester are highly correlated (0.90). Since total population is a critical predictor of disease spread, and it also appeared significant in the selection of teaching modes (Table 5), we kept total population in the model and dropped spring cases and deaths variables, and repeated the above analysis for robustness. The results are presented in the appendix Tables A1-A3. Treatment effect estimates in Table A1 and A2 are almost the same and corroborate our previous observations from the KMDM model. The noticeable difference in the logistic model results (Table A3) is that rural campus variable becomes statistically significant; i.e., colleges in rural areas are less likely to choose in-person over online. This can be attributed to the colleges in urban areas that decided to reopen with online modes due to

high COVID cases or deaths in the spring semester. Thus, dropping spring outcomes made the rural versus urban variable significant.

Table 3. Average treatment effects on COVID-19 cases
(Dependent variable: % change in cases compared to cases in two weeks before reopening date)

		(1)	(2)	(3)	(4)
Model	Treatment	0-15 Days	15-30 Days	30-45 Days	45-60 Days
1. Biv. OLS	1. In-person=1 online=0	18.291**	22.265	38.044	24.598
		(9.259)	(13.625)	(24.626)	(29.549)
	2. In-person=1 hybrid=0	12.768	3.537	27.221	21.753
		(10.273)	(20.489)	(30.055)	(34.139)
	3. Hybrid=1 online=0	5.552	18.532	22.661	19.474
		(8.497)	(16.705)	(17.007)	(23.512)
2. Mult. OLS	1. In-person=1 online=0	33.152**	39.162*	60.235	4.856
		(15.203)	(21.994)	(37.523)	(38.79)
	2. In-person=1 hybrid=0	23.629*	18.091	39.057	3.023
		(12.96)	(21.659)	(33.033)	(37.159)
	3. Hybrid=1 online=0	4.261	10.94	20.976	39.665
		(12.189)	(18.592)	(16.925)	(30.561)
3. PSM	1. In-person=1 online=0	35.052***	36.061*	49.915	21.246
		(13.174)	(19.77)	(33.179)	(33.564)
	2. In-person=1 hybrid=0	22.983*	17.79	41.661	15.955
		(13.02)	(22.032)	(33.546)	(39.944)
	3. Hybrid=1 online=0	6.607	3.782	18.706	26.166
		(10.74)	(16.098)	(19.398)	(44.361)
4. NN	1. In-person=1 online=0	24.459*	26.044*	46.144*	48.135*
		(12.881)	(14.811)	(25.293)	(27.656)
	2. In-person=1 hybrid=0	20.004**	18.501	41.576**	48.599*
		(9.765)	(14.026)	(20.643)	(28.815)
	3. Hybrid=1 online=0	7.311	11.901	10.299	15.661
		(11.304)	(18.141)	(13.583)	(21.643)
5. KMDM	1. In-person=1 online=0	35.014***	54.036***	79.443***	81.676**
		(10.631)	(15.375)	(28.319)	(36.443)
	2. In-person=1 hybrid=0	18.265	12.657	35.801	21.845
		(12.887)	(22.913)	(31.083)	(40.897)
	3. Hybrid=1 online=0	14.813	27.690	33.37**	53.353*
		(10.944)	(19.227)	(16.514)	(28.153)

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. N=1,076. Model 1 has no control variable. Models 2-5 have the same control variables. Control variables are listed in Table 5 and matching performance is reported in the appendix (Figure A1). State fixed effects were also used but not reported. Biv.=Bivariate, OLS=Ordinary Least Squares, Mult.=Multiple, PSM=Propensity Score Matching, NN=Nearest Neighbor, KMDM=Kernel Multivariate Distance Matching.

Table 4. Average treatment effects on COVID-19 deaths
(Dependent variable: % change in deaths compared to deaths in two weeks before reopening date)

Model	Treatment	(1) 0-15 Days	(2) 15-30 Days	(3) 30-45 Days	(4) 45-60 Days
1. Biv. OLS	1. In-person=1 online=0	6.957	10.057	10.966	1.841
		(8.163)	(10.674)	(11.535)	(15.012)
	2. In-person=1 hybrid=0	14.127	22.991**	13.781	17.915
		(8.809)	(11.044)	(12.559)	(17.359)
	3. Hybrid=1 online=0	3.847	0.669	9.568	-1.631
		(7.045)	(9.995)	(10.358)	(13.33)
2. Mult. OLS	1. In-person=1 online=0	13.789	23.266*	6.053	-15.816
		(11.488)	(13.931)	(14.993)	(22.799)
	2. In-person=1 hybrid=0	26.791**	23.111	12.67	22.044
		(11.299)	(15.512)	(14.917)	(21.657)
	3. Hybrid=1 online=0	-1.979	4.448	3.897	-17.015
		(9.098)	(12.649)	(13.085)	(16.629)
3. PSM	1. In-person=1 online=0	10.704	18.588	4.863	-2.444
		(10.877)	(14.882)	(16.026)	(26.242)
	2. In-person=1 hybrid=0	25.247**	9.418	-4.274	7.128
		(12.406)	(19.31)	(16.85)	(21.691)
	3. Hybrid=1 online=0	-1.216	-2.301	-1.16	-25.527
		(8.448)	(11.778)	(12.788)	(16.251)
4. NN	1. In-person=1 online=0	9.595	12.683	5.308	3.521
		(10.205)	(16.835)	(16.344)	(20.662)
	2. In-person=1 hybrid=0	11.968	8.134	12.879	41.827**
		(10.81)	(15.865)	(15.744)	(21.125)
	3. Hybrid=1 online=0	-10.882	-4.681	-6.098	-19.883
		(7.777)	(9.122)	(11.438)	(14.994)
5. KMDM	1. In-person=1 online=0	20.185**	29.215**	39.461***	38.404*
		(8.953)	(11.474)	(12.565)	(19.662)
	2. In-person=1 hybrid=0	24.329**	26.296*	24.879*	31.853
		(10.325)	(13.466)	(14.868)	(20.131)
	3. Hybrid=1 online=0	0.727	4.734	15.209	5.31
		(7.38)	(9.833)	(10.742)	(14.044)

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. N=1,076. Model 1 has no control variable. Models 2-5 have the same control variables. Control variables are listed in Table 5 and matching performance is reported in the appendix (Figure A1). State fixed effects were also used but not reported. Biv.=Bivariate, OLS=Ordinary Least Squares, Mult.=Multiple, PSM=Propensity Score Matching, NN=Nearest Neighbor, KMDM=Kernel Multivariate Distance Matching.

Table 5. Predictors of teaching modalities (Logistic regression coefficients)

Variables	(1) In-person=1 online=0	(2) Hybrid=1 online=0	(3) In-person=1 hybrid=0
% stayed home	-0.0289	-0.00890	0.0122
	(0.0514)	(0.0536)	(0.0380)
Log enrollment	-0.616***	-0.177	-0.476***
	(0.192)	(0.192)	(0.160)
Log cost of attendance	0.568	0.378	0.400
	(0.711)	(0.756)	(0.696)
Log endowment per student	-0.436***	-0.388***	-0.108
	(0.152)	(0.147)	(0.103)
Log student-faculty ratio	-0.820*	-1.125**	0.393
	(0.444)	(0.465)	(0.385)
Public=1, 0 otherwise	-0.895	-1.044	0.0880
	(0.666)	(0.703)	(0.625)
Log total population	-0.812***	-0.392	-0.155
	(0.289)	(0.287)	(0.239)
Black population (%)	0.00524	0.00325	0.0184
	(0.0203)	(0.0204)	(0.0170)
Male population (%)	-0.103	-0.0116	-0.0679
	(0.103)	(0.116)	(0.0870)
Log household income	1.338*	1.018	0.133
	(0.753)	(0.771)	(0.661)
Republican votes (%)	0.0462***	0.0156	0.0346**
	(0.0177)	(0.0178)	(0.0154)
Mask ordinance=1, 0 otherwise	1.731	-1.075	-0.141
	(1.126)	(1.337)	(1.254)
Rural campus=1, 0 otherwise	-0.761	-0.278	-0.186
	(0.478)	(0.516)	(0.372)
Log cases in spring 2020	0.334	0.224	0.00667
	(0.287)	(0.272)	(0.190)
Log deaths in spring 2020	-0.0715	-0.0367	-0.121
	(0.233)	(0.216)	(0.164)
Temperature in reopening month	0.0108	-0.0684	0.0940
	(0.0562)	(0.0557)	(0.0575)
County college count	0.0213	-0.0218	0.0430
	(0.0618)	(0.0356)	(0.0536)
Constant	1.868	1.836	-6.738
	(11.64)	(14.27)	(10.78)
Observations (colleges)	406+386=792	284+406=690	386+284=670
Accuracy	78.50%	70.45%	69.19%
Wald χ^2	149.68***	75.28**	78.33**
Log pseudolikelihood	-224.37	-243.49	-299.97
Pseudo R-squared	0.334	0.168	0.152

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. N=1,076. The third category is omitted from the sample for each pairwise comparison. For example, 406 online + 386 in-person colleges make 792 colleges in column (1), 284 hybrid and 406 online make 690 colleges in column (2), and so on. State fixed effects were added but not reported. We use log forms of some predictors to reduce the influence of extreme values on coefficient estimates. Accuracy shows the percentage of teaching modalities correctly identified by the model. The pseudo R-squared is higher (0.33) for the in-person versus online prediction (column 1) as opposed to hybrid versus online or in-person versus hybrid. This occurs because hybrid teaching modes are widely heterogeneous across colleges and harder to measure and categorize. Nevertheless, the main interest of this study is to analyze the difference between full in-person and full online, which also gives us directions on their mixture, i.e. hybrid modes. Moreover, pseudo R-squared cannot be interpreted as the goodness of fit because, unlike the R-squared from OLS, it shows the improvement in model likelihood over a null model (e.g., Heinzl et al. 2005; Hu et al. 2006; Hammert et al. 2018). Table 5 also presents the accuracy of the logistic models, suggesting that the teaching modalities of about seven to eight colleges out of 10 colleges were accurately predicted by our model.

5 Concluding remarks and implications

What do these results imply? Assume two similar counties with two comparable colleges, and each county reported 100 COVID cases before the fall 2020 semester. Further assume that the first county has its college reopened online, and has 10 more COVID cases in two months. If the college in the second county reopened in person, then its county may have 18 more cases in two months, ceteris paribus. Again, assume 100 COVID-related deaths before the fall 2020 semester in each of these two counties. If the first county observes three deaths in two months, the latter county observes four deaths. Moreover, if the second college reopens with a hybrid mode, then the number of cases and deaths respectively would be 15 and four in two months, because hybrid modes make a difference from in-person only at the beginning, but reaches up to the level of in-person within two months.

We found a statistically significant association between teaching modalities and COVID outcomes, which is consistent with several studies that find a negative impact of remote teaching or school closures on COVID outcomes (e.g., Auger et al. 2020; Pan et al. 2020; Andersen et al. 2020; Chernozhukov et al. 2021), but our result contradicts the studies that find little or no negative

impact (e.g., UNICEF 2020; Herby et al. 2022). However, these studies analyze teaching modalities or closures of schools, and not colleges where students are more likely to be affected due to their age and the frequency of socialization (e.g., Wrighton and Lawrence 2020; Goodman et al. 2021). The only exception is Andersen et al. (2020) who analyzed college reopenings. Their preliminary findings suggest that college reopenings increase COVID-19 cases by 1.7 daily cases per 100,000 residents in the first two weeks. Assuming there was no case pre-semester, the exponential growth of 1.7 cases per day for 60 days generates a large number. However, we control for many college and county-level factors that Andersen et al. (2020) did not include in their (preliminary) analysis, which might have contributed to the difference.

Our findings have several implications for higher education and health policy that simultaneously apply to agricultural economics and agribusiness departments and/or land grant universities. First, results from logistic regression suggest that colleges with bigger class sizes, having more populated campuses, or located in more populated areas chose remote teaching components, i.e., preferred online or hybrid to in-person modes. Although the rural campus variable is not a significant predictor in our results, many land-grant universities with urban or semi-urban campuses fall under this category. On the contrary, colleges with smaller endowments per student are less likely to choose online over in-person teaching modes, and hybrid over online modes, after controlling for other factors. Thus, relatively weaker positioned with heavy dependence on traditional on-campus tuition and auxiliary revenues would likely tend to return to the classroom with a quicker frequency out of financial exigency. Moreover, the average treatment effects show adopting an in-person mode instead of an online mode is associated with more COVID-19 cases. Therefore, colleges with small endowments need special policy attention to combat a disease-induced crisis.

Second, our results indicate that wealthier communities are more likely to have in-person teaching modalities, then less wealthy communities must have resources for online teaching. However, many studies discuss the added stress and the lack of social interactions in online instruction modes and hence poor student performance (e.g., Picault 2021; Lazarevic and Bentz 2021; Kofoed et al. 2021; Orlov et al. 2021). Thus, communities with low real income are in greater need of teaching resources on making online courses more interactive and performance-oriented.

Third, the analysis also suggests that campus reopening has a positive relationship with COVID-19 cases across all teaching modes, and both new cases and new deaths tend to increase for colleges teaching in-person compared to colleges teaching online. Since the treatment effects ideally find counterfactuals, one can point out that the expected risk of spreading a communicable disease can be partially mitigated with an initiative from the colleges by increasing distance education elements in classes. For hybrid modes, it is important to follow the containment policies at both the college and community level, as we found that the in-person components of hybrid modes may influence the disease spread and undermine the expected benefits from remote instruction.

Our study is observational and not based on a randomized placebo-controlled trial, hence should be interpreted with great caution. However, it still finds significant patterns in the data, offering important insights about choosing a teaching modality during a disease-induced pandemic, given the college and county-level features. Future research can look into the short-run and long-run effects of different teaching modalities on student learning outcomes.

References

1. Abadie, A., and G.W. Imbens. 2011. "Bias-corrected matching estimators for average treatment effects." *Journal of Business & Economic Statistics* 29(1):1-11.

2. ACS- American Community Survey. 2020. "American Community Survey Data." Retrieved from https://www.census.gov/programs-surveys/acs/data.html.

3. Andersen, M.S., A.I. Bento, A. Basu, C. Marsicano, and K. Simon. 2020. "College Openings, Mobility, and the Incidence of COVID-19 Cases." medRxiv. (Doi:10.1101/2020.09.22.20196048).

4. Atkeson, A. 2020. "What will be the economic impact of COVID-19 in the us? Rough estimates of disease scenarios." Working Paper 26867. National Bureau of Economic Research.

5. Auger, K.A., S.S. Shah, T. Richardson, D. Hartley, M. Hall, A. Warniment, K. Timmons et al. 2020. "Association between statewide school closure and COVID-19 incidence and mortality in the US." *Journal of American Medical Association* 324(9):859-870.

6. Austin, P.C. 2011. "An introduction to propensity score methods for reducing the effects of confounding in observational studies." *Multivariate Behavioral Research* 46(3):399-424.

7. Barnow, B.S., G.S. Cain, and A.S. Goldberger. 1981. "Issues in the Analysis of Selectivity Bias", In E.W. Stromsdorfer and G. Farkas, eds. *Evaluation Studies Review Annal 5*. Beverly Hills: Sage Publications, pp. 43-59.

8. Brahmasrene, T., and J.W. Lee. 2012. "Determinants of intent to continue using online learning: A tale of two universities." *Interdisciplinary Journal of Information, Knowledge, and Management* 7(1):1-20.

9. BTS -Bureau of Transportation Statistics. 2020. "Daily travel during the COVID-19 public health emergency." Retrieved from https://www.bts.gov/browse-statistical-products-and-data/trips-distance/daily-travel-during-covid-19- pandemic.

10. Chernozhukov, V., H. Kasahara, and P. Schrimpf. 2021. "The association of opening K–12 schools with the spread of COVID-19 in the United States: County-level panel data analysis." *Proceedings of the National Academy of Sciences* 118(42):e2103420118.

11. Cohen, J.A., D. Mistry, C.C. Kerr, and D.J. Klein. 2020. "Schools are not islands: Balancing COVID-19 risk and educational benefits using structural and temporal countermeasures." medRxiv. Doi:10.1101/2020.09.08.20190942.

12. De Paola, M., M. Ponzo, and V. Scoppa. 2013. "Class size effects on student achievement: heterogeneity across abilities and fields." *Education Economics* 21(2): 135-153.

13. Di Domenico, L., G. Pullano, G. Pullano, N. Hens, and V. Colizza. 2020. "Expected impact of school closure and telework to mitigate COVID-19 epidemic in France." Technical Report 8. France: Epicx Lab.

14. Fox, S.J., M. Lachmann, and L.A. Meyers. 2020. "Risks of COVID-19 introductions as schools reopen." COVID-19 Modeling Consortium, The University of Texas at Austin.

15. Gallagher, S., and J. Palmer. 2020. "The pandemic pushed universities online. The change was long overdue." *Harvard Business Review*, September. Retrieved from https://hbr.org/2020/09/the-pandemic-pushed-universities-online-the-change-was-long-overdue.

16. Gandolfi, A. 2020. "Planning of school teaching during COVID-19." *Physica D: Nonlinear Phenomena*, 415(January 2021):132753.

17. Goodman, K.E., L.S. Magder, J.D. Baghdadi, L. Pineles, A.R. Levine, E.N. Perencevich, and A.D. Harris. 2021. "Impact of sex and metabolic comorbidities on coronavirus disease 2019 (COVID-19) mortality risk across age groups: 66 646 inpatients across 613 US hospitals." *Clinical Infectious Diseases* 73(11):e4113-e4123.

18. HealthData.gov. 2020. "COVID-19 state and county policy orders." Retrieved from https://healthdata.gov/dataset/covid-19-state-and-county-policy-orders.

19. Heinzl, H., T. Waldhor, and M. Mittlbock. 2005. "Careful use of pseudo R-squared measures in epidemiological studies." *Statistics in Medicine* 24(18):2867-2872.

20. Hemmert, G.A.J., L.M. Schons, J. Wieseke, and H. Schimmelpfennig. 2018. "Log-likelihood-based pseudo R-squared in logistic regression: Deriving sample-sensitive benchmarks." *Sociological Methods and Research* 47(3):507-531.

21. Herby, J., L. Jonung, and S. Hanke. 2022. "A literature review and meta-analysis of the effects of lockdowns on COVID-19 mortality." *Studies in Applied Economics* 200(2022):1-62.

22. Hu, B., J. Shao, and M. Palta. 2006. "Pseudo-R-squared in logistic regression model." *Statistica Sinica* 16(2006):847-860.

23. Huber, M., M. Lechner, and A. Steinmayr. 2015. "Radius matching on the propensity score with bias adjustment: tuning parameters and finite sample behaviour." *Empirical Economics* 49(1): 1-31.

24. Hubler, S., and A. Hartocollis. 2020. "How colleges be- came the new COVID hot stops." The New York Times, September 11. Retrieved from https://www.nytimes.com/2020/09/11/us/college-campus-outbreak-covid.html.

25. Imbens, G.W. 2000. "The role of the propensity score in estimating dose- response functions." *Biometrika* 87(3): 706–710.

26. Jann, B. 2017. "Kernel matching with automatic bandwidth selection". London, UK: Stata. Retrieved from https://www.stata.com/meeting/uk17/slides/uk17_Jann.pdf.

27. Kim, H., A.M. Fox, Y. Kim, R. Kim, G. Bae, and M. Kang. 2021. "Is the male disadvantage real? Cross-national variations in sex gaps in COVID-19 incidence and mortality." *Global Public Health* 16(12):1793-1803.

28. King, G., and R. Nielsen. 2019. "Why propensity scores should not be used for matching." *Political Analysis*, 27(4):435-454.

29. Kofoed, M., L. Gebhart, D. Gilmore, and R. Moschitto. 2021. "Zooming to class?: Experimental evidence on college students & apos; online learning during Covid-19." IZA Discussion Paper 14356.

30. Lazarevic, B., and D. Bentz. 2021. "Student perception of stress in online and face-to-face learning: the exploration of stress determinants." *American Journal of Distance Education* 35(1): 2-15.

31. Leu, D.J., E. Forzani, and C. Kennedy. 2015. "Income inequality and the online reading gap: Teaching our way to success with online research and comprehension." *The Reading Teacher* 68(6):422-427.

32. Livadiotis, G. 2020. "Statistical analysis of the impact of environmental temperature on the exponential growth rate of cases infected by COVID-19". *PloS one* 15(5):e0233875.

33. McCaffrey, D.F, B.A. Griffin, D.A., M.E. Slaughter, R. Ramchand, and L.F. Burgette. 2013. "A tutorial on propensity score estimation for multiple treatments using generalized boosted models." *Statistics in medicine* 32(19):3388–3414.

34. McKibbin, W., and R. Fernando. 2020. "The economic impact of COVID-19." In R. Baldwin and W. di Mauro eds. *Economics in the Time of COVID-19*. London: CEPR Press. pp. 1-45.

35. Millett, G.A.., A.T. Jones, D. Benkeser, S. Baral, L. Mercer, C. Beyrer, B. Honermann. 2020. "Assessing differential impacts of COVID-19 on black communities." *Annals of epidemiology* 47(2020):37-44.

36. MIT Election Data and Science Lab. 2018. "County Presidential Election Returns 2000-2016" V6, Retrieved from the Harvard Dataverse. 10.7910/DVN/VOQCHQ.

37. NCES- National Center for Education Statistics. 2020. "Integrated Postsecondary Education Data System." U.S. Department of Education. Institute of Education Sciences. Retrieved from https://nces.ed.gov/ipeds/use-the-data.

38. NOAA- National Centers for Environmental Information. 2021. "County average temperature". Climate at a Glance: County Mapping. April. Retrieved from https://www.ncdc.noaa.gov/cag/county/mapping/50/tavg/202009/2/value.

39. NY Times- New York Times. 2020. "Coronavirus (COVID-19) Data in the United States." Retrieved from https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html.

40. Orlov, G., D. McKee, J. Berry, A. Boyle, T. DiCiccio, T. Ransom, A. Rees-Jones, and J. Stoye. 2021. "Learning during the COVID-19 pandemic: It is not who you teach, but how you teach." *Economics Letters* 202(2021):109812.

41. Pan, W.K., S. Tyrovolas, G. Iago, R.R. Dasgupta, F. Daniel, B. Zaitchik, P.M. Lantos, and C.W. Woods. 2020. "Covid-19: effectiveness of non-pharmaceutical interventions in the United States before phased removal of social distancing protections varies by region." medRxiv. Doi: https://doi.org/10.1101/2020.08.18.20177600.

42. Panovska-Griffiths, J., C.C. Kerr, R.M. Stuart, D. Mistry, D. J. Klein, R.M. Viner, and C. Bonell. 2020. "Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study." *The Lancet Child & Adolescent Health* 4(11):817–827.

43. Paul, S., and E. Lorin. 2021. "Distribution of incubation periods of COVID-19 in the Canadian context." *Scientific Reports* 11(1):1-9.

44. Picault, J. 2021. "Structure, flexibility and consistency: A dynamic learning approach for an online asynchronous course." *Applied Economics Teaching Resources* 3(4):30-43.

45. Pew Research Center. 2021. "Mobile Fact Sheet". April 7. Retrieved from pewresearch.org/internet/fact-sheet/mobile/

46. Rosenbaum, P.R., and D.B. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70(1):41–55.

47. Rosenbaum, P.R., and D.B. Rubin. 1984. "Reducing bias in observational studies using subclassification on the propensity score." *Journal of the American statistical Association* 79(387):516–524.

48. Siedner, M.J., G. Harling, Z. Reynolds, R.F. Gilbert, S. Haneuse, A.S. Venkataramani, and A.C. Tsai. 2020. "Social distancing to slow the US COVID-19 epidemic: Longitudinal pretest–posttest comparison group study." *PLoS medicine* 17(8):e1003244.

49. Sun, A., and X. Chen. 2016. "Online education and its effective practice: A research review." *Journal of Information Technology Education* 15(2016):157-190.

50. Tan, A.X., J.A. Hinman, H.S.A. Magid, L.M. Nelson, and M.C. Odden. 2021. "Association between income inequality and county-level COVID-19 cases and deaths in the US." *JAMA Network Open* 4(5):e218799-e218799.

51. Tyson, A. 2020. "Republicans remain far less likely than Democrats to view COVID-19 as a major threat to public health." Fact Tank by Pew Research Center, 1(1):1. Retrieved from https://www.pewresearch.org/fact- tank/2020/07/22/republicans-remain-far-less-likely-than-democrats-to-view- covid-19-as-a-major-threat-to-public-health/.

52. UNICEF- United Nation's Children's Fund. 2020. "In-Person schooling and COVID-19 transmission: A review of the evidence." November 23. Retrieved from

https://www.unicef.org/media/89046/file/In-person-schooling-and-covid-19transmission-review-of-evidence-2020.pdf.

53. Walke, H.T., M.A. Honein, and R.R. Redfield. 2020. "Preventing and responding to COVID-19 on college campuses." *Journal of American Medical Association* 324(17):1727-1728.

54. Wrighton, M.S., and S.J. Lawrence. 2020. "Reopening colleges and universities during the COVID-19 pandemic." *Annals of Internal Medicine* 173(8):664-665.

Appendix

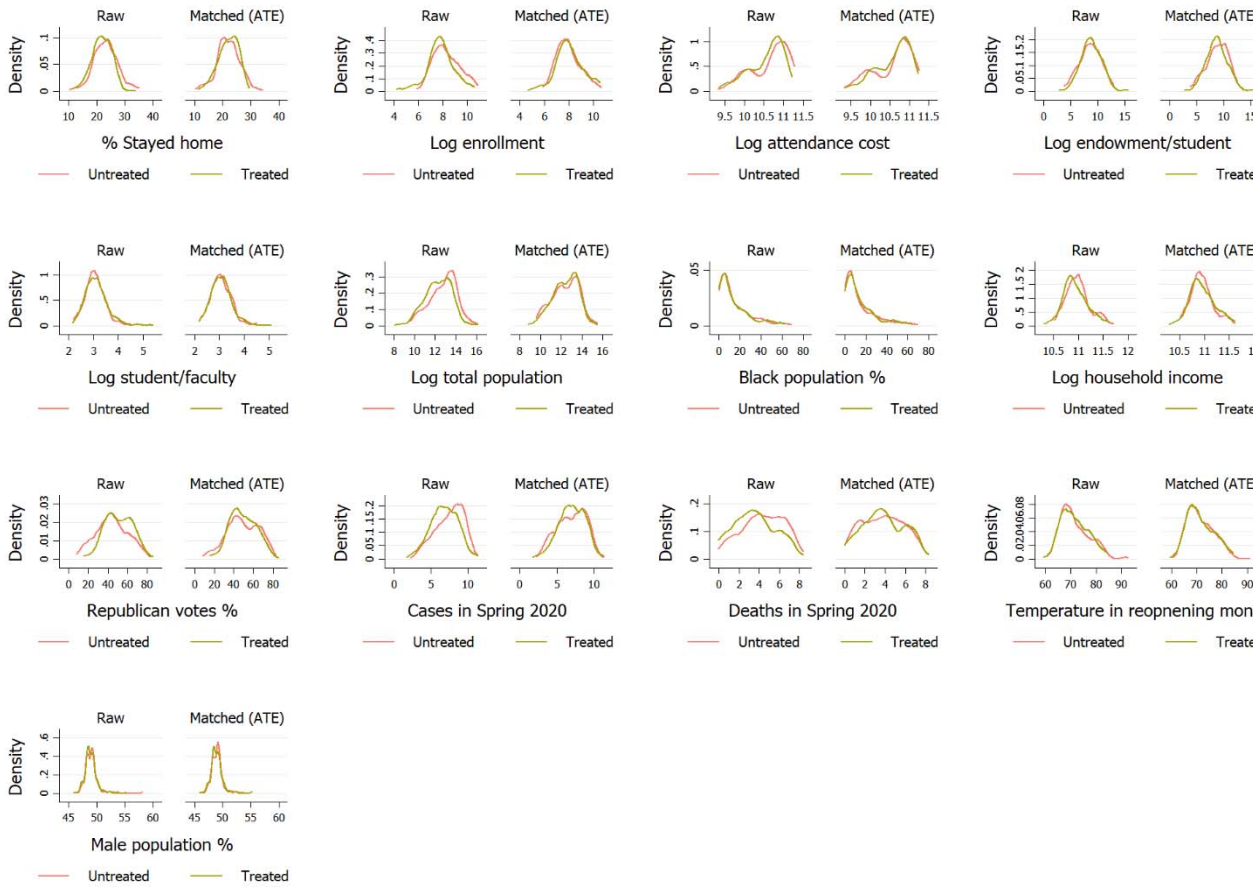

Figure A1. Densities of major variables before and after matching
Note: Figure shows kernel matching for multiple variables. See Table 1 for variable description and Table 2 for summary statistics.

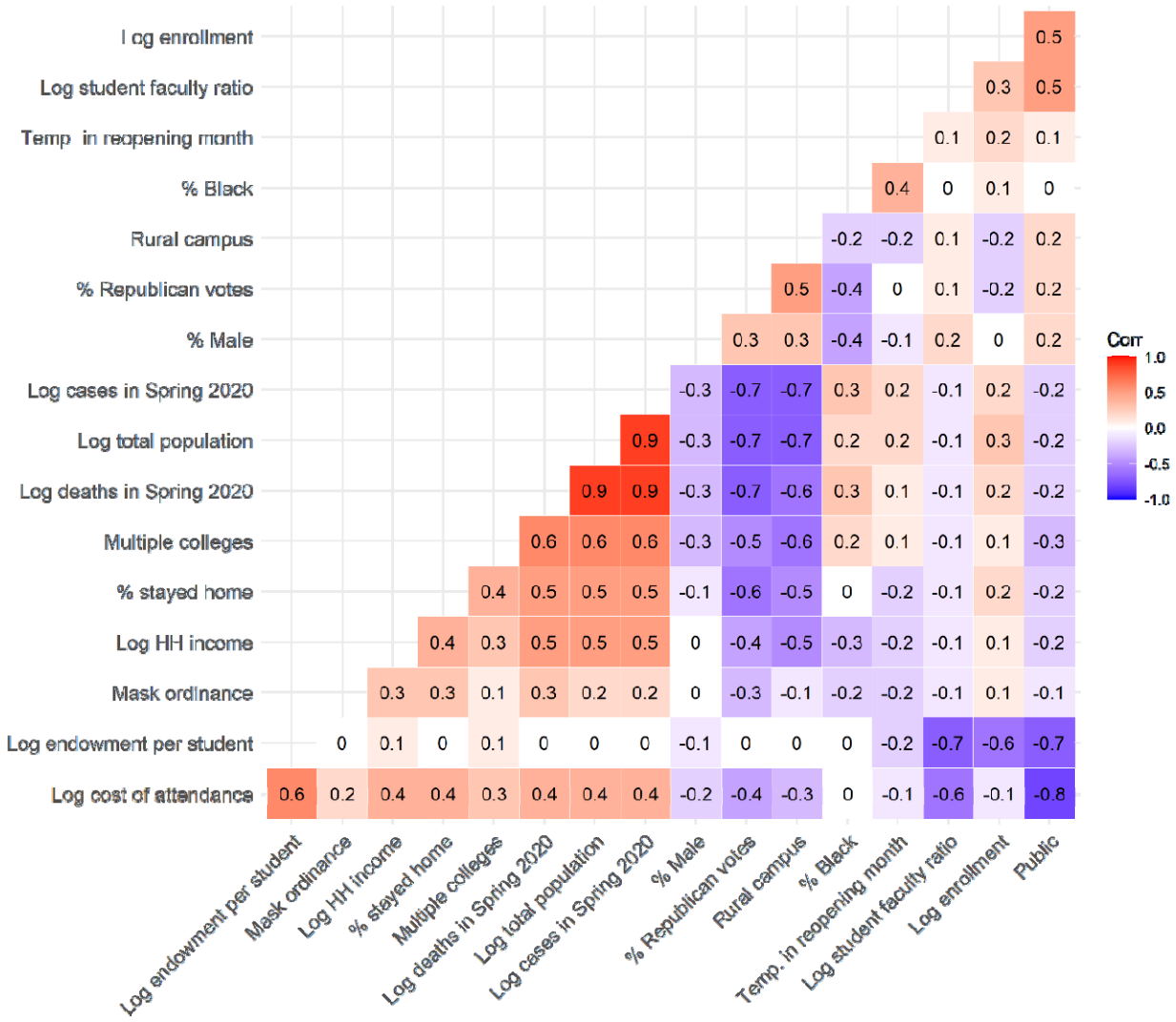

Figure A2. Correlation heatmap of control variables

Note: Figure shows Pearson correlation coefficients, rounded to one digit after the decimal and ordered by the strength of correlation. Some figures are too small and become zero when rounded. Darker shades mean greater association. See Table 1 for variable description and Table 2 for summary statistics.

Table A1. Average treatment effects on COVID-19 cases
(Dependent variable: % change in cases compared to cases in two weeks before reopening date)

		(1)	(2)	(3)	(4)
Model	Treatment	0-15 Days	15-30 Days	30-45 Days	45-60 Days
1. Biv. OLS	1. In-person=1 online=0	18.291**	22.265	38.044	24.598
		(9.259)	(13.625)	(24.626)	(29.549)
	2. In-person=1 hybrid=0	12.768	3.537	27.221	21.753
		(10.273)	(20.489)	(30.055)	(34.139)
	3. Hybrid=1 online=0	5.552	18.532	22.661	19.474
		(8.497)	(16.705)	(17.007)	(23.512)
2. Mult. OLS	1. In-person=1 online=0	31.273**	36.068*	54.87	1.804
		(14.957)	(20.95)	(35.865)	(37.521)
	2. In-person=1 hybrid=0	24.88*	21.364	43.458	4.768
		(12.798)	(21.711)	(34.555)	(37.635)
	3. Hybrid=1 online=0	3.229	9.693	18.837	36.561
		(12.012)	(18.04)	(16.673)	(31.07)
3. PSM	1. In-person=1 online=0	41.677***	41.11**	61.454	16.121
		(15.275)	(20.893)	(37.993)	(45.06)
	2. In-person=1 hybrid=0	27.596**	24.916	44.545	11.999
		(12.524)	(22.504)	(36.494)	(41.782)
	3. Hybrid=1 online=0	1.906	5.135	16.283	27.897
		(9.34)	(14.837)	(17.925)	(37.268)
4. NN	1. In-person=1 online=0	27.616**	29.32*	46.174	43.329
		(13.409)	(15.895)	(28.347)	(31.894)
	2. In-person=1 hybrid=0	23.906**	28.757	63.644	71.883
		(10.426)	(19.796)	(41.519)	(54.532)
	3. Hybrid=1 online=0	7.947	13.505	13.993	28.416
		(11.577)	(18.526)	(15.835)	(32.107)
5. KMDM	1. In-person=1 online=0	35.414***	55.546***	80.956***	83.528**
		(10.769)	(15.509)	(28.45)	(36.497)
	2. In-person=1 hybrid=0	18.504	13.74	38.934	25.924
		(13.133)	(23.634)	(31.916)	(41.911)
	3. Hybrid=1 online=0	14.418	27.957	33.382**	52.470*
		(11.016)	(19.9)	(16.667)	(29.037)

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. N=1,076. Model 1 has no control variable. Models 2-5 have the same control variables. Control variables are listed in Table 5 and matching performance is reported in the appendix (Figure A1). State fixed effects were also used but not reported. Biv.=Bivariate, OLS=Ordinary Least Squares, Mult.=Multiple, PSM=Propensity Score Matching, NN=Nearest Neighbor, KMDM=Kernel Multivariate Distance Matching.

Table A2. Average treatment effects on COVID-19 deaths
(Dependent variable: % change in deaths compared to deaths in two weeks before reopening date)

Model	Treatment	(1) 0-15 Days	(2) 15-30 Days	(3) 30-45 Days	(4) 45-60 Days
1. Biv. OLS	1. In-person=1 online=0	6.957	10.057	10.966	1.841
		(8.163)	(10.674)	(11.535)	(15.012)
	2. In-person=1 hybrid=0	14.127	22.991**	13.781	17.915
		(8.809)	(11.044)	(12.559)	(17.359)
	3. Hybrid=1 online=0	3.847	0.669	9.568	-1.631
		(7.045)	(9.995)	(10.358)	(13.33)
2. Mult. OLS	1. In-person=1 online=0	15.17	25.102*	9.92	-9.817
		(11.556)	(14.257)	(15.253)	(23.025)
	2. In-person=1 hybrid=0	26.529**	22.94	12.63	22.402
		(11.273)	(15.525)	(14.77)	(21.401)
	3. Hybrid=1 online=0	-1.445	3.825	4.825	-15.612
		(9.022)	(12.401)	(12.964)	(16.45)
3. PSM	1. In-person=1 online=0	20.741*	34.882**	15.899	8.536
		(10.993)	(14.927)	(15.411)	(25.666)
	2. In-person=1 hybrid=0	28.888**	19.464	2.484	20.774
		(12.643)	(19.031)	(20.004)	(24.016)
	3. Hybrid=1 online=0	-1.044	-4.021	2.707	-18.061
		(8.782)	(13.75)	(13.685)	(17.266)
4. NN	1. In-person=1 online=0	8.86	12.656	8.373	4.16
		(10.265)	(16.248)	(15.457)	(20.331)
	2. In-person=1 hybrid=0	16.503	14.129	14.104	42.81**
		(11.084)	(15.221)	(15.126)	(20.024)
	3. Hybrid=1 online=0	-10.836	-3.666	-2.027	-14.982
		(7.723)	(9.451)	(12.868)	(14.771)
5. KMDM	1. In-person=1 online=0	20.183**	28.705**	39.344***	37.87*
		(8.997)	(11.526)	(12.66)	(19.703)
	2. In-person=1 hybrid=0	23.417**	23.509*	23.593	27.159
		(10.21)	(13.772)	(15.323)	(21.276)
	3. Hybrid=1 online=0	0.507	4.235	15.202	4.996
		(7.369)	(9.814)	(10.844)	(14.394)

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. N=1,076. Model 1 has no control variable. Models 2-5 have the same control variables. Control variables are listed in Table 5 and matching performance is reported in the appendix (Figure A1). State fixed effects were also used but not reported. Biv.=Bivariate, OLS=Ordinary Least Squares, Mult.=Multiple, PSM=Propensity Score Matching, NN=Nearest Neighbor, KMDM=Kernel Multivariate Distance Matching.

Table A3. Predictors of teaching modalities (Logistic regression coefficients)

Variables	(1) In-person=1 online=0	(2) Hybrid=1 online=0	(3) In-person=1 hybrid=0
% stayed home	-0.035	-0.008	0.009
	(0.047)	(0.053)	(0.038)
Log enrollment	-0.612***	-0.192	-0.459***
	(0.186)	(0.190)	(0.158)
Log cost of attendance	0.516	0.386	0.413
	(0.715)	(0.752)	(0.692)
Log endowment per student	-0.420***	-0.400***	-0.099
	(0.145)	(0.144)	(0.103)
Log student-faculty ratio	-0.758*	-1.130**	0.401
	(0.438)	(0.462)	(0.382)
Public=1, 0 otherwise	-0.894	-1.067	0.092
	(0.661)	(0.699)	(0.623)
Log total population	-0.522**	-0.183	-0.283
	(0.227)	(0.196)	(0.176)
Black population (%)	0.009	0.006	0.016
	(0.0197)	(0.019)	(0.016)
Male population (%)	-0.104	-0.004	-0.069
	(0.0997)	(0.116)	(0.087)
Log household income	1.565**	1.195	0.110
	(0.749)	(0.744)	(0.655)
Republican votes (%)	0.0434**	0.014	0.0348**
	(0.018)	(0.017)	(0.015)
Mask ordinance=1, 0 otherwise	1.680	-0.933	-0.232
	(1.099)	(1.332)	(1.240)
Rural campus=1, 0 otherwise	-0.794*	-0.280	-0.190
	(0.465)	(0.507)	(0.371)
Temperature in reopening month	0.037	-0.058	0.082
	(0.056)	(0.052)	(0.055)
County college count	0.027	-0.018	0.039
	(0.061)	(0.035)	(0.053)
Constant	-3.388	-2.391	-4.542
	(11.410)	(12.910)	(10.280)
Observations (colleges)	406+386=792	284+406=690	386+284=670
Accuracy	78.38%	70.21%	67.89%
Wald χ^2	140.87***	74.16**	77.57**
Log pseudolikelihood	-225.72	-244.05	-300.49
Pseudo R-squared	0.332	0.166	0.152

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. N=1,076. The third category is omitted from the sample for each pairwise comparison. For example, 406 online + 386 in-person colleges make 792 colleges in column (1). State fixed effects were added but not reported. We use log forms of some predictors to reduce the influence of extreme values on coefficient estimates. Accuracy shows the percentage of teaching modalities correctly identified by the model.