Variation in the timing of Covid-19 communication across universities in the UK

Alejandro Quiroz Flores[1*], Farhana Liza[1¶], Husam Quteineh[1¶], Barbara Czarnecka[2¶]

[1] Business and Local Government Data Research Centre, University of Essex, Colchester, Essex, United Kingdom

[2] Division of Management, Marketing and People, Business School, London South Bank University, London, United Kingdom

* Corresponding author

E-mail: aquiro@essex.ac.uk (AQF)

¶ These authors contributed equally to this work

Abstract

During the Covid-19 pandemic, universities in the UK used social media to raise awareness and provide guidance and advice about the disease to students and staff. We explain why some universities used social media to communicate with stakeholders sooner than others. To do so, we identified the date of the first Covid-19 related tweet posted by each university in the country and used survival models to estimate the effect of university-specific characteristics on the timing of these messages. In order to confirm our results, we supplemented our analysis with a study of the introduction of coronavirus-related university webpages. We find that universities with large numbers of students are more likely to use social media and the web to speak about the pandemic sooner than institutions with fewer students. Universities with large financial resources are also more likely to tweet sooner, but they do not introduce Covid-19 webpages faster than other universities. We also find evidence of a strong process of emulation, whereby universities are more likely to post a coronavirus-related tweet or webpage if other universities have already done so.

Introduction

University responses to the spread of respiratory illnesses

Pandemic outbreaks of respiratory illnesses have struck universities for hundreds of years. European universities have documented the effect of pandemics since at least the fourteenth century [1-2]. Many American universities closed their campuses during the influenza pandemics of the twentieth century [3-7]. More recently, universities in Asia were severely affected by the 2009 H1N1 pandemic and the 2002-04 SARS outbreak.

Universities have incentives to prepare and respond to outbreaks of respiratory illnesses because they affect student health, reduce academic performance, and lead to increased use of health care [4, 6, 8-13].

In this context, universities' first line of defence is influenza vaccination. While seasonal vaccination does not protect against the uncommon viruses at the heart of pandemics, they provide a basic level of protection [10] and reduce visits to doctors and health centres, as well as reduce hospitalisation. As a second line of defence against outbreaks of respiratory illnesses, universities implement non-pharmaceutical interventions, including isolation, social distancing, smothering of coughs and sneezes, washing hands, and cleaning touched objects and surfaces, among others [14-18].

Regardless of the specific interventions implemented to mitigate outbreaks of respiratory illnesses, universities must rely on timely and effective communication campaigns. In this light, we present a study of the timing of university communication during the height of the Covid-19 pandemic in the UK.

University communication and social media during the Covid-19 pandemic: a crisis informatics approach to studying the impact of the pandemic on higher education

In early March 2020, Covid-19 had spread across the UK. At that time, the central government had not issued university-specific advice. Therefore, universities activated their response systems and implemented their own measures to control the disease on their campuses. In the first stage, universities raised awareness, reinforced public health advice, and provided guidance to students and staff [19]. Later on, they implemented more stringent measures, including social distancing and remote working for staff, particularly in mid-March 2020 when preparations for a national lockdown were in progress. In spite of some initial hesitation, universities closed their campuses to non-essential services by 23rd March 2020. This variation in university responses to Covid-19 motivated us to look more closely at how universities reacted to the pandemic.

The initial information campaign on university campuses and the subsequent implementation of interventions were announced to students and staff through email and internal newsletters. These emails and newsletters are private tools of internal crisis communication and the research team did not have systematic access to them. Yet, part of this engagement was observable in universities' social media channels, as universities are aware that students may prefer social media posts rather than emails [20], and this provided us with a unique opportunity to study how universities responded to the pandemic. Our investigation indicates that UK universities were making references to Covid-19 in social media since late January 2020. These social media posts generally raised awareness, reinforced public health advice, and provided guidance.

The public has been using social media and other forms of communication during crises to learn and inform themselves [21-22]. Organisations have embraced social media to enable rapid interaction with stakeholders [23-26]. Universities also use social media to communicate with students and staff in a frequent, timely, open, and targeted manner [27-32].

The use of social media as a two-way communication channel between universities and students and staff during the Covid-19 pandemic places our research in the area of *crisis informatics* [33-39]. Crisis informatics is a relatively new field that explores the role of information and communication technology (ICT) in crises. Specifically, it focuses on how networked ICT facilitates the public's response to a crisis. The field covers different types of crises, although it is particularly useful for the study of exogenous events such as natural hazards [37].

As the role of social media has become more important during crises, crisis informatics has made significant advances in several subjects, including the role of networked ICT on socio-behavioural factors during emergencies and the use of digital communication as a data source [37, 39, 40]. At the same time, there are challenges emerging from very large quantities of unstructured, noisy information. However, if the appropriate methods are applied to the collection, pre-processing, and analysis of data, social media can provide useful information for empirical analysis [37-39, 41-42].

We rely on crisis informatics to contribute to the emerging research agenda on the impact of Covid-19 on higher education [43]. This research agenda, while fragmented and microscopic [44-45], is making important contributions to our understanding of the effects of the SARS-CoV-2 virus and the pandemic on higher education. Currently, the emphasis has been on the disruption to traditional learning and the transition to online learning [46-50], as well as on the challenges in this transition, particularly for universities in developing countries [51-52, 43].

Research has also been devoted to the timing and heterogeneity of non-pharmaceutical interventions during the height of the pandemic [53-54]. Closely linked to this strand of work are epidemiological simulations for university campuses that inform university interventions, including contact tracing and quarantining [55-56]. Interventions are supported by communication efforts and recent research has focused on communication strategies [57-60, 20, 45, 51], particularly on the use of social media and its positive effects on student satisfaction with university responses to the crisis [20, 45].

The pandemic not only affected students but also university staff, both physically and in terms of additional work pressure and general uncertainty. Thus, recent research has focused on the mental and physical health of staff [61-62], and the key role of social support [61]. Recent work is also addressing the role of university leadership in managing the effects of the pandemic on campuses around the world and new studies are confirming the positive effect of women in managing the crisis [63, 20].

In summation, this paper explores the timing of coronavirus-related messages posted by universities in social media. Research shows that the timing of interventions can reduce the negative effects of pandemic outbreaks [64]. This is particularly pertinent to risk communication and therefore *our aim is to explain why some universities posted social media messages sooner than others*. In order to confirm our results, we supplemented our analysis of social media with a study of the introduction of coronavirus-related university webpages, which were also widely used by universities to communicate Covid-19 information to stakeholders [53].

Theoretical framework

In order to explain variation in the timing of communication, we rely on Situational Crisis Communication Theory (SCCT) and theories of policy emulation.

During crises, organisations engage in strategic communication. According to Situational Crisis Communication Theory (65-66), institutions have strong incentives to communicate early with stakeholders when they are also victims of a crisis. This is often the case when natural disasters, including pandemics, take place–stakeholders do not attribute the crisis to the organisation, which in turn can benefit from providing information about the emergency. In fact, research evidence suggests that early communication by an organisation when a crisis is attributed to external factors contributes to the perceived credibility of the organisation (65-69).

This logic is particularly important for UK universities in the context of the pandemic. According to SCCT, UK higher education institutions are victims of the pandemic and this gives them incentives to provide early information to their stakeholders in order to gain credibility. Institutional credibility was crucial because UK universities had to compete for students in the highly uncertain admission cycle of 2020. In this context of urgency and competition, our empirical analysis focuses on the variables that best reflect universities' organisational capacity and ability to communicate early with students and staff.

Theories of policy emulation also help us understand the variation in the timing of university communications. While there are nuances across theories of emulation, they generally focus on the opportunities for policy diffusion: "Policy diffusion is the process whereby a state is more likely to adopt a policy if other states have already adopted that policy." [70] We follow this literature and focus on the role of geographic proximity as a source of diffusion, which is best exemplified by Tobler's first 'law' of geography where "everything is related to everything else, but near things are more related than distant things." [71] More recent research adds a second 'law:' "Everything resembles everything else, but closer things are more similar" [71]. In terms of crisis communication, we expect that universities are more likely to communicate early with their stakeholders if universities in their vicinity have already done so.

In sum, our study contributes to our understanding of risk communication in the higher education sector during the pandemic and to our knowledge of the implementation of non-pharmaceutical interventions across campuses in the UK. These interventions, and the communication efforts that support them, are important because they slow down the spread of infection on campuses, thus reducing the negative effects of the pandemic on student health, academic performance, and use of health care. Moreover, and in the context of the pandemic in the UK, universities filled a vacuum caused by the absence of central government advice to higher education institutions. In so doing, universities were confirming their key role as public sources of trust and potentially reducing the negative effects of a decline of the higher education sector in the UK economy. Universities, as victims of the crisis, quickly engaged their stakeholders and raised awareness, reinforced public health advice, and provided guidance through social media, in order to meet their duty of care and gain credibility in an uncertain admissions cycle.

Material and methods

In order to explain why some universities posted Covid-19-related social media messages sooner than others, we followed a two-fold strategy. First, we collected posts and their metadata from universities' official Twitter accounts to identify the date of their first Covid-19-related tweet during the height of the Covid-19 pandemic. Second, we used these dates to estimate Cox survival models of elapsed time and survival models of diffusion to explore the role of emulation. We used these two types of models to explore whether universities choose the timing of communication based only on their university-specific characteristics or whether they also considered actions taken by other institutions.

To test the validity of our findings from Twitter data, we applied the research design described above to the dates of universities' first official Covid-19 webpages.

Twitter data

The crisis informatics literature explores several peer-to-peer communication platforms [35]. A large proportion of the research focuses on social media, including "blogging and microblogging, social networking sites, social media sharing platforms, and wikis" [42]. Although universities use multiple social media platforms, we focus on Twitter because most UK universities have a Twitter account. In addition, Twitter's emphasis on text, as well as the wide availability of computational methods to pre-process Twitter content and analyse text as data, make it a suitable source of information for the analysis of risk communication. In this sub-section we describe how we identified universities' first tweet with Covid-19 content.

As a first step, we focused on the Twitter accounts of all officially recognised universities and colleges in the UK as higher learning institutions that can award degrees [72]. This list includes 170 universities, although our sample consists of 166 universities because some institutions do not have a Twitter account, while others have ceased operations or their business model is mainly online teaching, which was not as severely affected by the pandemic. We manually reviewed the Twitter accounts used in this paper to confirm their authenticity. In addition, we replicated our analyses of Table 1 using only accounts verified by Twitter; these results are presented in S1 Table. Twitter verifies accounts that are determined to be in the public interest; this assures the public that these Twitter profiles are authentic.

As a second step, we collected tweets posted between 31st December 2019 –when the WHO first identified a statement from Wuhan Municipal Health Commission related to a new 'viral pneumonia'– and the end of our study on 6th May 2020. We used Twitter's public API to collect tweets that provided data encoded in JavaScript Object Notation (JSON). The extraction produced 57,340 tweets for the 166 universities in our sample within our period of interest.

We focus on two attributes of tweets: the text content and the timestamp. The content of a tweet may contain non-textual characters, including URLs, mentions, hashtags, emojis, or numbers. We used text pre-processing methodologies to improve the quality of the data, mitigate the creative use of spacing and punctuation, and remove non-textual content. These methodologies include separating hyperlinks from the adjacent text, normalising Twitter-specific tokens (e.g., hashtags and URLs), extracting text from in between symbols, replacing ampersands, lowercasing the text, normalising multiple occurrences of vowels and consonants, normalising emojis and numbers, splitting numbers and emojis when adjacent to text, and removing non-alphanumeric characters. In general, we used text normalisation to produce text concordant with standard natural language processing approaches applied to formal text.

Once we pre-processed all tweets, we applied tokenisation to obtain a bag of words from each tweet. We then applied pattern-matching rules to extract tweets that mention the pandemic. Specifically, we used four keywords: ‘coronavirus’, ‘covid’, ‘COVID-19’, and ‘face-to-face.’ Our initial search had a more extensive set of keywords for pattern matching, but it produced a large set of irrelevant tweets. After some manual exploration, we found that these four keywords captured the most relevant tweets for the study; they are also a better reflection of the strict measures that universities would eventually implement, including the end of face-to-face teaching.

These pre-processing and tokenisation methods reduced our original sample of 57,340 tweets to 7,015 relevant tweets. We then simply ranked them by timestamp to select the first tweet of each university. We manually cross-checked the first tweet for each university and removed any results that produced a tweet that was not relevant to our search. Thus, our final sample includes the date of the first Covid-19 related tweet for 158 universities.

University-specific characteristics

We use survival analysis –also known as hazard analysis or event history modelling– to analyse why some universities posted Covid-19 related tweets sooner than others. This method focuses on time to an event or a transition. In biostatistics, for example, the emphasis may be on a patient’s time to death or remission after a cancer diagnosis [73]. In this paper, our event of interest is the first Covid-19 related tweet posted by a university. Thus, the dependent variable (*Days to Tweet*) is the number of days from 31st December 2019 to the date of a university’s first Covid-19 related tweet. In our sample of 158 universities, 153 posted a Covid-19 related tweet; the remaining five universities did not post a first tweet by the end of our study and therefore we coded them as right-censored. Our data indicates that the median time to posting the first tweet is 66 days with a 95 per cent confidence interval of 62 to 72 days.

Fig 1 presents a more systematic analysis of the number of days to post the first tweet about Covid-19. The figure presents the Kaplan-Meier estimate of the survival function, which in this case can be interpreted as the proportion of universities that *have not posted* a Covid-19 tweet over time. On 31st December 2019, not a single university had mentioned the novel coronavirus, but as time went by, more and more institutions posted a tweet about it. By 23rd March, almost all universities in the UK had mentioned something about Covid-19 at least once.

Fig 1. Survivor function of days to first Covid-19 tweet.

There seem to be three periods in this graph. The first period is between 31st December 2019 and 24th January, when few universities posted their first tweet. In the second period, starting at the end of January, a larger number of universities posted their first Covid-19 related tweet, thus reducing the survival function drastically–by 28th February 2020, when the first internal transmission was recorded in the UK, about 45 per cent of universities had already posted their first message. The third period starts in early March, when the preparations for strict non-pharmaceutical interventions were underway–by 13th March, about 70 per cent of universities had posted their first tweet. By 23rd March, almost all universities had posted at least one Covid-19 message on Twitter.

Why did some universities tweet sooner than others? In this section, we explore if universities choose the timing of their first tweet based *only* on their university-specific characteristics, such as the size of the student population or university financial resources. To the best of our knowledge, our paper presents the first analysis of the role of university-specific characteristics on the timing of communication during the Covid-19 pandemic.

First, we expect that universities with larger numbers of students will post a Covid-19 tweet sooner than universities with fewer students. We conjecture that most students and staff received official university messages about the pandemic over email or internal newsletters, but that there is a proportion of individuals who would not read those messages. For universities with a large number of students, that proportion could equate to thousands of individuals. In this case, posting messages and announcements in Twitter and other social media channels might be an effective way of reaching out to students and staff–messages are short and to the point, and can be re-posted by peers and colleagues, thus potentially reaching the students and staff who may not have read internal communications. In this case, posting a tweet sooner rather than later can be an effective way to raise awareness of the pandemic and provide guidance and advice to students and staff.

To measure the size of the student population in universities, we obtained the total number of student enrolments by higher education provider and applied a natural logarithm transformation to this number to produce the variable (*ln(Total Enrolment*)). This logarithmic transformation represents the orders of magnitude of student numbers and allows us to compare cases where some universities have more than 40,000 students and others have fewer than 500. We do not control for staff numbers because they are highly correlated with the size of the student population, thus creating a collinearity problem.

Our second set of expectations is related to resilience. The largest effect of the pandemic on UK universities will be caused by a decrease in student numbers [74]. In this light, our baseline model (Model 1) controls for additional university-specific factors that make universities more or less resilient to a negative shock to student numbers.

Our first control variable is the proportion of university income dependent on tuition fees. We expect that universities that rely heavily on tuition fees are more sensitive to a negative shock in student numbers than universities that are more research-oriented. The proportion of income dependent on tuition fees, which we label (*Proportion Income Tuition*), is simply the ratio of tuition fees to total income. Total income is composed of tuition fees, funding body grants, research grants, investment income, donations, and other income.

Our second control variable is university total reserves. Reserves are a measure of wealth and we expect that wealthy universities have the necessary resources to protect students and staff, and the capacity to endure a drastic reduction in student numbers. Total reserves are measured in millions of pounds sterling and include all types of university reserves, both restricted and unrestricted. As with the number of student enrolments, we applied a natural logarithm transformation to this variable to account for a large variation in the data; we labelled this variable (*ln(Total Reserves*)).

We excluded the University of Oxford and the University of Cambridge from all our analyses because they have financial resources that are incomparable to the resources of other universities, even when a logarithmic transformation is applied. Excluding Cambridge and Oxford, the mean total reserves for our sample of universities is £218 million. In contrast, Cambridge has £5.1 billion in total reserves while Oxford has £4.1 billion in reserves. We also removed from the analysis a very small number of universities that had negative total reserves.

Our third control variable is interaction with the public. This variable is measured as the number of attendants to free events, including lectures, performances, exhibitions, museums, and other events. As with the total number of student enrolments, we applied a natural logarithm transformation to account for a large variation in the data; we labelled this variable (*ln(Public Interaction)*). Interaction with the public is a double-edged sword, as it may increase the risk of infection through exposure but also strengthen resilience in terms of links to the community.

Our fourth control variable indicates whether a university is a member of the Russell Group of universities: (*Russell Group*). This variable is equal to one if a university is one of the 24 universities in the Russell Group and equal to zero otherwise. We expect that universities in this group will be more resilient because they are older –which provides experience in dealing with crises– but also because they have large financial resources and are research-intensive, which allows them to endure negative shocks to student numbers. We obtained the list of Russell Group universities from the group's official website.

S2 Table presents additional analyses that control for the gender of university vice-chancellors and for the proportion of positions in university leadership teams occupied by women. As mentioned in the introduction, the characteristics of the leadership of an organisation play an important role on crisis response [75-76, 20], and recent work on Covid-19 indicates that women are more effective in reducing Covid-19 deaths [63]. Results from S2 Table indicate that the gender of university vice-chancellors and the proportion of positions in university leadership teams occupied by women do not have a statistically significant effect on the timing of communication. The names and gender of university vice-chancellors were obtained from Universities UK [77] and from official university websites.

The proportion of positions in university leadership teams occupied by women were obtained from official university websites.

To summarise, our baseline Model 1 of university-specific characteristics includes the *log(Total Enrolment*), *Proportion Income Tuition*, *Total Reserves*, *ln(Public Engagement)*, and *Russell Group* membership. In addition, we estimated two alternative models. Model 2 includes a measure of campus size as given by the number of university buildings per number of students and staff (*Buildings per capita*). Model 3 replaces *ln(Total Reserves*) with *ln(Unrestricted Reserves*). Unrestricted reserves, measured in millions of pounds sterling, are a component of total reserves but do not include sensitive sources of funds, such as a university's endowment. We note again that we eliminated Oxford and Cambridge from all our analyses due to their enormous financial resources–the mean unrestricted reserves for our sample is £157 million. In contrast, Cambridge has over £3 billion in unrestricted reserves while Oxford has £2.8 billion. We also eliminated a handful of universities with negative unrestricted reserves.

The variables ln(Total Enrolment), Proportion Income Tuition, ln(Total Reserves), ln(Public Engagement), Buildings per capita, and ln(Unrestricted Reserves), were obtained from the Higher Education Statistics Authority (HESA) [78]. These variables correspond to the academic year 2018-19, with the exception of the number of buildings, which corresponds to the academic year 2017-18. These were the most recent statistics available from HESA when we completed our study and we believe that they have not changed drastically for the academic year 2019-20. Thus, they continue to provide an adequate reflection of university-specific characteristics during the height of the pandemic. Summary statistics for all variables for the estimation sample of our baseline Model 1 in Table 1 are presented in S3

Table. The specific tables from HESA used to support the findings of our study are presented in S4 Appendix.

We now turn to our estimation procedure. Table 1 presents three Cox-semiparametric models of our dependent variable *Days to Tweet*, which is the number of days from 31st December 2019 to the date of a university's first Covid-19 related tweet. All models in this paper were estimated in Stata version 15. We use Cox models because we do not have a strong theory about the shape of the hazard rate and therefore we prefer to leave it unparametrized. As long as the proportionality assumption is met by the models, this choice does not affect the substantive effects of our variables of interest.

We applied four different specifications of proportional hazards tests available in Stata 15 to all Cox models in this paper, including analysis time, the log of analysis time, one minus the Kaplan-Meier product-limit estimate, and the rank of analysis time [79]. All models passed either all four tests or at least two of them; we are confident that they meet the proportionality assumption. The tests are available in our replication files. If a model passed only two tests out of four, we decided not to adjust the non-proportional covariate because all variables in our Cox models are time-invariant and the proper solution to the problem is unlikely to bring large benefits while causing drastic changes to the research design [80].

The estimation results in Table 1 consist of hazard ratios –that is, exponentiated coefficients– and their standard errors clustered for the upper-tier local authority (UTLA) to address a potential lack of independence for universities within the same authority. An UTLA is a geographic unit in the UK often identical to a county, unitary authority, or London borough. For ease of interpretation of Table 1, a hazard ratio above one indicates an increase in the hazard rate–this is the rate at which universities post their first tweet over time since 31st December 2019. In contrast, a hazard ratio below one indicates a decrease in the hazard rate. As an illustration, a hazard ratio of 1.3 indicates that a change in a covariate increases the hazard rate in 30 per cent, while a hazard ratio of 0.8 indicates a decrease of 20 per cent.

Table 1: Cox Models of Days to First Covid-19 Tweet.

	Model 1	Model 2	Model 3
Ln(Total Enrolment)	1.387***	1.397**	1.486***
	(0.153)	(0.188)	(0.189)
Proportion Income Tuition	0.487	0.474	0.365**
	(0.256)	(0.273)	(0.179)
Ln(Total Reserves)	1.294**	1.302**	
	(0.162)	(0.174)	
Ln(Public Interaction)	0.883**	0.879**	0.897*
	(0.0505)	(0.0540)	(0.0520)
Russell Group	1.424	1.451	1.563
	(0.514)	(0.559)	(0.562)
Buildings per capita		0.000148	
		(0.00150)	
Ln(Unrestricted Reserves)			1.131
			(0.143)
Observations	141	135	139
Subjects	141	135	139
Failures	139	133	137
Clusters	88	87	87
Log L	-550.7	-520.9	-542.3

Dependent variable: Days to first Covid-19 tweet. Event of interest: First Covid-19 tweet. Results in hazard ratios. Standard errors in parentheses clustered on UTLA. Oxford, Cambridge, and universities with negative total and negative unrestricted reserves are excluded from the analyses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Emulation

In this section we investigate if universities consider the actions of other institutions in their decision to post a first Covid-19 related tweet. To do so, we estimate survival models of emulation used in the literature on public policy diffusion [81-88]. As mentioned in the introduction, “Policy diffusion is the process whereby a state is more likely to adopt a policy if other states have already adopted that policy.” [70]

We do not aim to understand the causes of emulation –which may be connected to competition, for instance– but to look for evidence of a diffusion process across UK universities. To the best of our knowledge, this is the first study of diffusion in university communication during the Covid-19 pandemic.

Recent models of diffusion rely on dyadic data whereby pairs of states or countries are the unit of statistical analysis [83-84, 89]. We follow this literature and use dyads of UK universities as units of analysis. For example, we create the dyad Essex-Bristol, Essex-Kent, Essex-Roehampton, and so on. For 170 universities, there are 170^2=28,900 university dyads. Each of these dyads is followed daily from 31st December 2019 to 6th May 2020, which gives us a potential sample of 3,670,300 observations. Our sample is smaller because many universities posted their first tweet before the 6th of May.

This daily dyadic setup for our data is useful because we can record the date when a university tweets for the first time and track if other universities have tweeted before in order to explore the likelihood of emulation. It is precisely for this reason that the dyad Essex-Kent is not the same as the dyad Kent-Essex: Kent may emulate Essex if Essex tweeted first, but Essex cannot emulate Kent.

In the daily dyad University A-University B, our dependent variable (*Emulation*) is equal to one on the day when University A posts its first tweet if University B has previously posted a tweet, and zero otherwise. The literature on diffusion prescribes that once *Emulation* takes on a value of one on a particular date, it should then be coded as missing; this is simply because we focus on time to emulation and because once two universities have taken the same course of action, that is, posting a tweet, emulation is no longer a possibility. For the estimation sample of Model 2 in Table 2, there are 5,831 cases where *Emulation* is equal to one and 853,141 cases where it is equal to zero.

In a dyad, University A cannot emulate University B if the latter has not posted a tweet in the first place. Thus, our key determinant of emulation is a variable labelled (*B Tweeted*) that is equal to one if University B has tweeted and equal to zero otherwise. For example, in the dyad University A-University B, the former might tweet on 10$^{\text{th}}$ March while the latter tweeted on 5$^{\text{th}}$ March. For the estimation sample of Model 2 in Table 2, there are 171,382 cases where *B Tweeted* is equal to one and 687,590 cases where it is equal to zero. As prescribed in the literature on diffusion, the variable *B Tweeted* would then be equal to zero from 31$^{\text{st}}$ December 2019 to 4$^{\text{th}}$ of March and equal to one from 5$^{\text{th}}$ March onwards. We do not expect that a tweet will have an immediate effect and therefore we use a two-day lag of this event. Our results are robust to the use of three and four-day lags for *B Tweeted*; S5 Table presents these additional estimation results.

Our dyadic data has all 28,900 dyads and a university may emulate any other university. Nevertheless, we expect that universities may be more responsive to the actions of their geographical neighbours because they share similar infection risks. Thus, we created a variable (*Neighbour*) that indicates whether two universities in a dyad are geographical neighbours. The variable *Neighbour* is equal to one if two universities are separated by a distance of 50 kilometres of less, and equal to zero otherwise. For the estimation sample of Model 2 in Table 2, approximately 12 per cent of universities are neighbours according to this definition.

In S6 Table, we present estimates from the dyadic models of Table 2 using two alternative definitions of geographic proximity. In the first alternative, two universities are neighbours if they are separated by a distance of 100 kilometres of less. In the second alternative, two universities are neighbours if they are separated by a distance of 25 kilometres of less. Our results are robust to these alternative definitions of a neighbourhood.

To calculate distances between universities, we used the Google Maps API to request the full address of each university, including its longitude and latitude. We then used these coordinates and the package 'geodist' [90] in *R* version 3.5.0 to create a matrix of distances for each dyad. We follow the literature on diffusion described above and interact the variable *Neighbour* with the variable *B Tweeted*. This interaction of variables allows us to investigate whether the likelihood of emulation depends on geographic proximity.

In addition to testing for the presence of diffusion, we use our research design to analyse the effect of the daily number of coronavirus infections in a university's upper tier local authority. We focused on infection cases rather than deaths because the recording of Covid-19 related deaths in England is still a matter of debate. The number of Covid-19 cases was obtained from Public Health England as reported in Coronavirus (COVID-19) in the UK [91]. For the estimation sample of Model 2 in Table 2, the mean daily number of Covid-19 cases is 0.3 with a variance of 3.32; the minimum number is zero and the maximum is 33. We also applied a natural logarithm transformation to the number of Covid-19 cases and created the variable (*ln(Covid-19 Daily Cases)*). We collected this data on 30th April 2020 and therefore estimation is restricted to days between 31st December 2019 and 30th April 2020. This does not affect our analyses, as most universities had posted their first tweet by the end of March 2020.

The cases of Covid-19 are reported at the upper-tier local authority (UTLA) level in England. Unfortunately, these figures are not reported for Wales, Scotland, and Northern Ireland. However, we were able to include observations from universities in Wales, Scotland, and Northern Ireland until 30 January 2020, when there were no reported cases of infections in the UK. There are efforts to collect and organise coronavirus cases for Scotland and Wales using medical wards [92], but these are not comparable to the UTLAs in England.

We matched universities to UTLAs using their coordinates as explained above and assigning them to the polygons of UTLAs. These polygons were obtained from the Office of National Statistics file on Counties and Unitary Authorities (December 2017) Full Clipped Boundaries in UK [93]. We used the package 'sp' [94] in R version 3.5.0 to assign university coordinates to UTLA polygons.

Our models of diffusion also control for all the university-specific variables used in the previous section. Although these variables do not change between 31st December 2019 and 6th May 2020, they are useful indicators of university-specific characteristics. Our controls include university total reserves, and therefore we exclude Oxford, Cambridge, and universities with negative total reserves from our analyses of emulation.

In dyadic models, it is also recommended that specifications include control variables for both University A and University B [84]. This is simply because the probability of emulation depends on the actions of the two universities: the leader and the follower. Thus, all specifications include controls for both universities in a dyad, which we separate with subscripts. For instance, Model 2 in Table 2 controls for the natural logarithm of total student enrolments in University A, denoted, Ln(Total Enrolment)$_A$, and for the natural logarithm of total student enrolments in University B, denoted, Ln(Total Enrolment)$_B$.

We note that the literature on diffusion finds that traditional dyadic models create a bias in favour of an emulation effect. The intuition behind the bias is as follows: "Simply put, state *i* appears to emulate state *j* not because it looks to state *j* as a policy leader, but because both are independently headed in the same direction and state *j* may just happened to get there first." [84] In other words, the traditional dyadic model cannot distinguish if variables increase the likelihood that University B will implement a policy (and therefore that there is an opportunity for emulation) or if they increase the probability that University A will emulate University B. The solution to this bias is quite simple; rather than estimating the original, unconditional dyadic model, one needs to estimate a model that conditions on a university's opportunity to emulate. In this light, the purpose of the conditional model is not to find evidence of emulation but to distinguish if specific variables have an effect on emulation or on coincidental convergence.

In practical terms, in the conditional dyadic setup, the dependent variable is also *Emulation*, but the estimation sample is restricted to those days when there is an opportunity for emulation, that is, those days after University B has posted its first tweet. Thus, we condition on the variable (*Opportunity*), which is equal to one if University B has tweeted and equal to zero otherwise. For the estimation sample of Model 2 in Table 2, there are 163,670 cases where *Opportunity* is equal to one and 695,302 cases where it is equal to zero. In the dyadic conditional model where estimation is restricted to the 163,670 cases where *Opportunity* is equal to one, there are 5,831 cases where *Emulation* is equal to one and 157,839 cases where it is equal to zero. We note that the variable Opportunity is not identical to the variable B Tweeted because the opportunity to emulate starts the day after University B has tweeted.

Table 2 presents three models: a monadic survival model of universities' first tweet, a dyadic unconditional model of emulation, and a conditional model of emulation. The goal of the first model is to explore the effect of Covid-19 cases on the hazard rate of posting a first Covid-19 related tweet. The previous section did not explore the effect of infections simply because it uses a cross-section of universities, while the data for infections is measured daily. Thus, it was more appropriate to present this test here because it uses the same daily data organisation than the dyadic models. Having said this, the purpose of Model 2 in Table 2 is to look for evidence of emulation. Model 3 is the conditional model of emulation and its goal is to differentiate the effect of variables in the likelihood of emulation or coincidental convergence.

Lastly, we note that all models in Table 2 are discrete survival models [95-96]. Discrete survival models are implemented as models for binary choice –in our case, a logit model– that controls for duration dependence by adding a cubic polynomial of days between 31st December 2019 and the event of interest [96]. In our case, the event of interest in Model 1 is a university's first tweet, while in Models 2 and 3 the event of interest is emulation. Results for all models are presented in odds ratios. Standard errors clustered at University A in the dyad University A-University B are presented in parentheses in order to account for a potential lack of independence among observations. As in the previous section, we excluded the University of Oxford and the University of Cambridge, as well as any universities with negative total or unrestricted reserves.

Table 2: Models of first Covid-19 tweet and emulation of first Covid-19 tweet.

	Model 1: First Covid-19 tweet (monadic)	Model 2: Emulation of first Covid-19 tweet (dyadic unconditional)	Model 3: Emulation of first Covid-19 tweet (dyadic conditional)
Ln(Total Enrolment)$_A$	1.440***	1.536***	1.514***
	(0.188)	(0.212)	(0.211)
Proportion Income Tuition $_A$	0.750	0.211	0.235
	(0.649)	(0.204)	(0.229)
Ln(Total Reserves) $_A$	1.328*	1.065	1.093
	(0.194)	(0.145)	(0.154)
Ln(Public Interaction) $_A$	0.796**	0.919	0.915
	(0.0707)	(0.0682)	(0.0708)

Russell Group $_{A}$	1.921	1.172	1.168
	(0.763)	(0.564)	(0.569)
Ln(Covid-19 Daily Cases) $_{A}$	2.341***	1.505***	1.375**
	(0.390)	(0.239)	(0.215)
Days $_{A}$	1.137***	0.828***	0.562***
	(0.0512)	(0.0363)	(0.0373)
Days$^{2}{}_{A}$	0.998*	1.005***	1.011***
	(0.000966)	(0.00109)	(0.00146)
Days$^{3}{}_{A}$	1.000**	1.000***	1.000***
	(0.00000575)	(0.00000717)	(0.00000915)
Ln(Total Enrolment)$_{B}$		0.955***	0.980
		(0.0109)	(0.0131)
Proportion Income Tuition $_{B}$		1.226**	2.243***
		(0.114)	(0.274)
Ln(Total Reserves) $_{B}$		1.018	0.963*
		(0.0175)	(0.0190)
Ln(Public Interaction) $_{B}$		1.009	1.024***
		(0.00670)	(0.00718)
Russell Group $_{B}$		1.014	1.228***
		(0.0303)	(0.0489)
Ln(Covid-19 Daily Cases) $_{B}$		1.413***	1.362***
		(0.0757)	(0.0653)
B Twitted $_{A(t-2)}$		27.91***	
		(7.157)	
(Neighbour)(B Twitted $_{A(t-2)}$)		0.646***	
		(0.0545)	
Neighbour			0.682***
			(0.0537)
Constant	0.000102***	0.000132***	3.957
	(0.000109)	(0.000126)	(4.227)
Observations	6930	858972	163670
Clusters	131	141	140
Pseudo-R2	0.167	0.352	0.175
Log L	-480.9	-22637.1	-20753.8

Dependent variable (Model 1): First Covid-19 tweet. Dependent variable (Models 2-3): Emulation of first Covid-19 tweet. All models are discrete survival models with logit link and cubic polynomial for number of days to event. Results in odds ratios. Standard errors in parentheses clustered by university A. Oxford, Cambridge, and universities with negative total reserves are excluded from the analyses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Webpages data

In this section we supplement our analysis of Twitter data with information from university webpages. Universities also used Covid-19 dedicated webpages to raise awareness of the pandemic and provide guidance and advice to students and staff [53]. We acknowledge that the content of Covid-19 specific webpages is different than Twitter posts–webpages require more careful planning and implementation than tweets, as well as constant updating and maintenance. It is precisely for this reason that an analysis of webpages is important, as any confirmation of substantive results will give more confidence to the analysis presented in the previous section.

Our research design is the same as in our analysis of Twitter data. We explored if universities introduce Covid-19 webpages based on their own factors *and* the actions taken by other universities. We also used the same estimation methods. First, we used Cox models for the analysis of the number of days to posting a first webpage, as well as the same university-specific control variables. Second, we used models of diffusion and controlled for the same time-varying variables as in the previous section, including the introduction of webpages by other universities and the number of Covid-19 cases in a university's UTLA.

We began by identifying the date when universities first introduced a webpage with Covid-19 related information. We first mapped UK universities to their corresponding web domains, for instance essex.ac.uk. We then used the Google Search API to search every domain from 31st December 2019 to 6th May 2020 for the Covid-19 related keywords: 'Covid-19', 'Corona,' and 'Coronavirus.' The returned results for each matching page included a summary snippet, a title, and a Uniform Resource Locator (URL).

Unlike tweets, these webpages are as noisy as they are heterogenous in design, and a one-size-fits-all approach to noise reduction would not be useful to extract content. Therefore, our text extraction was limited to the page body, which allowed us to focus on the main text in a webpage while limiting noise in the navigation menus or announcements that contain Covid-19 related terms. This process produced 13,265 matching webpages for 128 universities.

We sorted these matching webpages by date and manually inspected the top result for each university to minimise noise. We used the dates from these webpages to produce the dependent variable (*Days to Webpage*), which is the number of days from 31st December 2019 to the date of a university's first Covid-19 webpage as described above. We note that we do not have any right-censored cases because our data collection produced a sample of 128 universities with a webpage. Unfortunately, we cannot be sure that the remaining 42 universities in the UK did not introduce a Covid-19 webpage and therefore it would be incorrect to code them as right-censored. Having said this, our data indicates that the median time to posting the first webpage is 55 days with a 95 per cent confidence interval of 43 to 63 days. Fig 2 presents the Kaplan-Meier estimate of the survival function of the number of days to introduce a webpage about Covid-19.

Fig 2. Survivor Function of Days to First Covid-19 Webpage.

We now turn to our estimation strategy. For the Cox models, our dependent variable is the number of days from 31st December 2019 to the date when a university first introduced a Covid-19 webpage.

Table 3 presents three Cox-semiparametric models of our dependent variable (*Days to Webpage*). As in the previous section, we use Cox models because we do not have a strong theory about the shape of the hazard rate and therefore we prefer to leave it unparametrized. Likewise, the estimation results in Table 3 consist of hazard ratios with their standard errors clustered for the upper-tier local authority (UTLA) presented in parentheses.

Table 3: Cox Models of Days to First Covid-19 Webpage.

	Model 1	Model 2	Model 3
Ln(Total Enrolment)	1.340***	1.482***	1.361***
	(0.150)	(0.220)	(0.153)
Proportion Income Tuition	0.188***	0.254***	0.175***
	(0.0999)	(0.128)	(0.0923)
Ln(Total Reserves)	1.058	1.079	
	(0.139)	(0.154)	
Ln(Public Interaction)	0.982	1.004	0.987
	(0.0525)	(0.0506)	(0.0526)
Russell Group	1.134	1.027	1.286
	(0.429)	(0.395)	(0.474)
Buildings per capita		271972.2	
		(2475205.5)	
Ln(Unrestricted Reserves)			0.980
			(0.108)
Observations	111	106	109
Subjects	111	106	109
Failures	111	106	109
Clusters	77	76	76
Log L	-409.0	-384.7	-400.0

Dependent variable: Days to first Covid-19 webpage. Results in hazard ratios. Standard errors in parentheses clustered on UTLA. Oxford, Cambridge, and universities with negative total and unrestricted reserves are excluded from the analyses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We now turn to our analysis of emulation, which used the same specifications as the emulation models of Twitter data, although the key determinant of emulation in this section is a variable labelled (*B Webpage*) that is equal to one if University B has introduced a Covid-19 webpage and equal to zero otherwise. For the estimation sample of Model 2 in Table 4, there are 99,214 cases where *B Webpage* is equal to one and 307,954 cases where it is equal to zero.

We estimated three models: one monadic model of universities' first Covid-19 related webpage, and two dyadic models of emulation, one unconditional and one conditional. For the estimation sample of Model 2 in Table 4, there are 3,405 cases where *Emulation* is equal to one and 403,763 cases where it is equal to zero. In the same sample, there are 95,459 cases where *Opportunity* is equal to one and 311,709 cases where it is equal to zero. Conditioning the analysis to the 95,459 cases where *Opportunity* is equal to one, there are 3,405 cases where *Emulation* is equal to one and 92,054 cases where it is equal to zero.

Table 4 presents results in odds ratios, which reflect changes in the odds of posting a first Covid-19 related webpage in Model 1 and the odds of emulation in Models 2-3. Standard errors clustered at University A in dyad University A-University B are presented in parentheses in order to account for a potential lack of independence among observations. As in the previous section, we excluded the University of Oxford and the University of Cambridge, as well as any universities with negative total or unrestricted reserves.

Our results are robust to the use of three and four-day lags for *B Webpage* for Model 2 in Table 4 (results presented in S7 Table), and to alternative definitions of a neighbourhood in the dyadic models of Table 4 (results presented in S8 Table).

Table 4: Models of first Covid-19 webpage and emulation of first Covid-19 webpage.

	Model 1: First Covid-19 webpage	Model 2:	Model 3:

	(monadic)	Emulation of first Covid-19 webpage (dyadic unconditional)	Emulation of first Covid-19 webpage (dyadic conditional)
Ln(Total Enrolment)$_A$	0.952	1.277	1.276
	(0.145)	(0.218)	(0.217)
Proportion Income Tuition $_A$	0.203*	0.677	0.678
	(0.185)	(0.774)	(0.774)
Ln(Total Reserves) $_A$	1.338*	1.207	1.211
	(0.208)	(0.231)	(0.232)
Ln(Public Interaction) $_A$	1.015	1.004	1.004
	(0.0565)	(0.0526)	(0.0524)
Russell Group $_A$	0.862	0.724	0.721
	(0.537)	(0.347)	(0.349)
Ln(Covid-19 Daily Cases) $_A$	1.921**	1.616*	1.600*
	(0.612)	(0.402)	(0.396)
Days $_A$	1.203**	1.102	1.034
	(0.101)	(0.0694)	(0.0670)
$\text{Days}^2{}_A$	0.997*	0.998	0.999
	(0.00185)	(0.00136)	(0.00138)
$\text{Days}^3{}_A$	1.000*	1.000	1.000
	(0.0000123)	(0.00000850)	(0.00000856)
Ln(Total Enrolment)$_B$		0.996	1.004
		(0.0115)	(0.0109)
Proportion Income Tuition $_B$		0.882***	0.931*
		(0.0428)	(0.0373)
Ln(Total Reserves) $_B$		1.023**	1.006
		(0.00932)	(0.00961)
Ln(Public Interaction) $_B$		0.989**	0.989**
		(0.00410)	(0.00424)
Russell Group $_B$		1.009	0.984
		(0.0134)	(0.00968)
Ln(Covid-19 Daily Cases) $_B$		1.076	1.070
		(0.0617)	(0.0605)
B Webpage $_{A(t-2)}$		30.79***	
		(4.353)	
(Neighbour)(B Webpage $_{A(t-2)}$)		0.837**	
		(0.0748)	

Neighbour			0.849*
			(0.0725)
Constant	0.000691***	0.00000388***	0.000426***
	(0.00111)	(0.00000558)	(0.000625)
Observations	4061	407168	95459
Clusters	81	109	109
Pseudo-R2	0.136	0.327	0.127
Log L	-312.4	-13252.3	-12821.1

Dependent variable (Model 1): First Covid-19 webpage. Dependent variable (Models 2-3): Emulation of first Covid-19 webpage. All models are discrete survival models with logit link and cubic polynomial for number of days to event. Results in odds ratios. Standard errors in parentheses clustered by university A. Oxford, Cambridge, and universities with negative total reserves are excluded from the analyses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Results and discussion

We organise our discussion around our two sets of results. First, we discuss the effects of the size of the student community, the role of universities' financial resources, and the impact of Covid-19 infections on the hazard rate of posting a first Covid-19 related tweet. We then consider if evidence from our analysis of university webpages supports our conclusions. Second, we discuss the role of emulation and whether estimation results are consistent across our two sources of data.

In order to guide our discussion, we focus on the hazard ratios of independent variables, and particularly if they are above one (increase hazard rates) or below one (decrease hazard rates), at an alpha level of 0.05. For consistency, we apply the same terminology to Cox models and our discrete survival models with logits links.

University size, financial wealth, and infections

Our first expectation is related to the size of the student community. We conjectured that not all university students and staff read university Covid-19 announcements communicated via email and internal newsletters. In fact, students may prefer social media posts rather than emails [20]. Therefore, universities have incentives to reinforce these announcements through social media; these incentives are stronger in large institutions simply because the number of individuals who may not have read private messages is larger. Thus, we expected that the number of student enrolments would increase the hazard rate of posting a first Covid-19 tweet. The hazard ratios for Ln(Total Enrolment) in the Cox models of Table 1 and the monadic model of Table 2 are well above one and statistically significant. This indicates that changes to the natural logarithm of total enrolment –which can also be interpreted as the elasticity of enrolment or per cent changes in total enrolment– increase the hazard rate of posting a tweet. *In other words, universities with larger numbers of students tweeted sooner than universities with fewer students*. This effect is also present in our analyses of universities with verified Twitter accounts, presented in S1 Table.

Our second set of expectations focuses on university-specific characteristics that determine resilience to a negative shock in student numbers. While there are multiple characteristics that deserve discussion, we highlight the role of financial resources because we expect that they will increase university resilience in the same way that countries' wealth strengthens disaster preparedness and response [97-99].

The hazard ratios for Ln(Total Reserves) in the Cox models of Table 1 are well above one and statistically significant, which indicates that per cent changes in total reserves increase the hazard rate of posting a Covid-19 related tweet. While the monadic model of Table 2 indicates that the hazard ratio for Ln(Total Reserves) is significant only at an alpha level of 0.1, our analyses of universities with verified Twitter accounts in Table 2 confirm that wealth increases the hazard rate. *Altogether, we find that wealthier universities were more likely to tweet sooner than universities with more modest means*.

In the context of the current pandemic, we explored the effect of the number of Covid-19 cases on the hazard rate of posting a Covid-19 related tweet. To do so, we used the daily number of infections in universities' upper-tier local authority as a control variable in our monadic model of universities' first Covid-19 related tweet in Table 2. The hazard ratio for Ln(Covid-19 Daily Cases)$_A$ in this model is well above two and statistically significant, which indicates that *per cent changes in Covid-19 cases greatly increase the hazard rate of posting a first Covid-19 tweet.*

We also note that the hazard ratio for Ln(Covid-19 Daily Cases)$_A$ in the dyadic models of Table 2 are also above one and significant, which indicates that Covid-19 infections also increase the likelihood of emulation. The fact that the coefficients for Ln(Covid-19 Daily Cases)$_A$ are quite similar across the unconditional and conditional dyadic models suggests that infections are driving emulation and not coincidental convergence.

We now consider if the effects of the size of the student community, the role of universities' financial resources, and the impact of Covid-19 infections in our Twitter data are also present in our analyses of university webpages.

We acknowledged that data from Twitter can be quite noisy and therefore we supplemented our analyses with information from official university Covid-19 webpages. We identified the date when universities first introduced a Covid-19 webpage and then applied the same research design implemented for our Twitter data to estimate survival models and models of diffusion. To summarise these results, *the analyses from webpages provide moderate support for the effect of university size and indicate that university financial resources do not have a statistically significant effect on the hazard rate of introducing a webpage. Nonetheless, these analyses confirm the effect of Covid-19 infections on the odds of introducing a first Covid-19 webpage*.

First, the hazard ratios for Ln(Total Enrolment) in the Cox models of Table 3 are well above one and statistically significant, which indicates that per cent changes in student enrolments increase the hazard rate of introducing a Covid-19 webpage. However, the monadic model of a first Covid-19 webpage in Table 4 indicates that student enrolments do not have a significant effect on the rate of introducing a webpage. We consider that this is only moderate support for the effect of the size of the student community on the hazard rate of introducing a webpage.

Moreover, the models do not find support for an effect of university financial resources. In fact, all Cox models in Table 3 find that Ln(Total Reserves) does not have a statistically significant effect, while the monadic model of a first Covid-19 webpage in Table 4 indicates that university resources would increase the hazard rate of introducing a webpage only at an alpha level of 0.1. This suggests that university reserves do not determine the likelihood of introducing a Covid-19 webpage.

Nevertheless, our analyses of webpage data confirm the effect of Covid-19 infections on the timing of risk communication. Indeed, the hazard ratio for Ln(Covid-19 Daily Cases)$_A$ in the monadic model of universities' first webpage in Table 4 is well above one and statistically significant, which indicates that per cent changes in Covid-19 cases increase the hazard rate of tweeting. We also observed this effect in our analysis of universities' first Covid-19 tweet.

It is also important to note that the hazard ratio for Ln(Covid-19 Daily Cases)$_A$ in the dyadic models of Table 4 is also above one and significant, which indicates that Covid-19 infections also increase the likelihood of emulation. As with Twitter data, the coefficients for Ln(Covid-19 Daily Cases)$_A$ are very similar across the unconditional and conditional dyadic models, which suggests that infections are driving emulation and not coincidental convergence.

Emulation

One of the central features of our research design is the estimation of models of diffusion. We estimated conditional and unconditional dyadic models of emulation to explore whether universities choose the timing of communication based only on their own university-specific characteristics or whether the actions of other universities also contributed to their response. As mentioned, we do not aim to understand the causes of emulation but to look for evidence of a diffusion process across UK universities.

Our unconditional dyadic model of emulation in Table 2 indicates that the hazard ratio for B Tweeted $_{A(t-2)}$ is very well above one and statistically significant. This suggests that universities are much more likely to follow institutions that have previously posted a Covid-19 related tweet. This effect is also present when we use three and four-day lags for B Tweeted, as indicated in our supplementary analyses in S5 Table. Evidence of emulation is one of the strongest results in our analyses and it is also replicated in our study of university webpages.

Interestingly, while a follower's likelihood of emulation is higher when other universities have posted a tweet, this likelihood is not as high if the leading university is a geographical neighbour, as demonstrated by the hazard ratio for (Neighbour)(B Tweeted $_{A(t-2)}$) in Table 2, which is smaller than one and statistically significant. We confirmed this effect in our supplementary analyses in S6 Table, which use two alternative definitions of a neighbourhood.

Our analyses of university webpages strongly confirm that universities are more likely to emulate if other institutions have previously posted a Covid-19 related webpage. Results from Table 4 indicate that the hazard ratio for B Webpage $_{A(t-2)}$ is also very well above one and statistically significant. The results are of the same magnitude, direction, and significance as in our analyses of Twitter data– this is a very strong indication of the effect of diffusion in university responses during the pandemic. Moreover, this effect is also present when we use three and four-day lags for B Webpage, as indicated in our supplementary analyses in S7 Table. They also confirm that while a follower's likelihood of emulation is higher when other universities have posted a webpage, this likelihood is not as high if the leading institution is a geographical neighbour, even when different definitions for a neighbourhood are used for estimation, as demonstrated in S8 Table.

These results point to a form of inequality among universities in the UK. Our estimation results indicate that universities with large student communities are quicker to engage in risk communication as measured by the timing of their first Covid-19 tweet and their first Covid-19 webpage. While all universities have similar incentives to reach out sooner to larger numbers of students during crises, the ability to do so depends on wealth. It is therefore not a coincidence that our estimation results suggest that universities with large financial resources, as measured by total reserves, are also quicker to engage in risk communication over social media.

Universities with large student communities and vast financial resources have something else in common: age. In the UK, a university's age is crucial because it brings wealth and experience with previous crises, and research shows that this has a positive effect in prevention [100, 97]. This simply means that older universities are wealthier, larger, and more experienced, and altogether more resilient to pandemics. These characteristics allow them to engage in risk communication at an earlier stage than other universities. Smaller, poorer, younger universities are not so resilient and this is reflected in the timing of their risk communication, which lags behind the efforts of more established universities. This coincides with the finding by the Institute of Fiscal Studies that universities with weak financial positions before the pandemic are at higher risk of insolvency as a result of the shock to student numbers [74].

On the more positive side, our analyses show that universities learn from each other. This means that there is a space for leadership and an opportunity for coordination during crises. While some coordination was organised by Universities UK, in terms of the negative consequences of the pandemic on universities' financial positions, there is a need for better coordination in the delivery of risk communication and the sharing of best practice that can allow the system to learn more quickly and respond more effectively to crises.

Indeed, a more effective crisis response would reduce the negative effects of the pandemic on the education sector and its link to the national economy. The UK education sector produces close to six per cent of national output and in the second quarter of 2020 it was estimated that 90 per cent of this output would be lost due to the pandemic [101]. At that time, multiple studies predicted that UK universities would lose billions of pounds in the long run and that some institutions would not be financially viable without significant government assistance [101, 74]. Our study shows that UK universities engaged in swift crisis communication in the absence of central government guidelines, which probably reduced some of the negative consequences of the pandemic.

In this light, we draw important lessons for universities around the world and contribute to our general understanding of the effects of the pandemic on higher education. Although the empirical results are only valid for institutions in the UK, the paper provides a useful research design that can be replicated for data on university responses in other countries

[53, 59]. In addition, our theoretical framework and selection of covariates, as well as the emphasis on survival analysis and models of policy diffusion, will serve as useful guidelines for further research.

References

1. Courtenay WJ. The effect of the Black Death on English higher education. Speculum. 1980 Oct 1;55(4):696-714.
2. Junker D. Six Hundred Years: Heidelberg University-Crisis and Achievements. Reviews of infectious diseases. 1987 Sep 1;9(Supplement_5):S439-42.
3. Markel H, Lipman HB, Navarro JA, Sloan A, Michalsen JR, Stern AM, Cetron MS. Nonpharmaceutical interventions implemented by US cities during the 1918-1919 influenza pandemic. Jama. 2007 Aug 8;298(6):644-54.
4. Layde PM, Engelberg AL, Dobbs HI, Curtis AC, Craven RB, Graitcer PL, Sedmak GV, Erickson JD, Noble GR. Outbreak of influenza A/USSR/77 at Marquette University. Journal of Infectious Diseases. 1980 Sep 1;142(3):347-52.
5. PONS VG, CANTER J, DOLIN R. Influenza A/USSR/77 (H1N1) on a university campus. American Journal of Epidemiology. 1980 Jan 1;111(1):23-30.
6. Sobal J, Loveland FC. Infectious disease in a total institution: a study of the influenza epidemic of 1978 on a college campus. Public health reports. 1982 Jan;97(1):66.
7. Smelser NJ. Reflections on the University of California: from the Free Speech Movement to the global university. Univ of California Press; 2010 Mar 4.
8. Nichol KL, Heilly SD, Ehlinger E. Colds and influenza-like illnesses in university students: impact on health, academic and work performance, and health care use. Clinical Infectious Diseases. 2005 May 1;40(9):1263-70.
9. Nichol KL, D'Heilly S, Ehlinger E. Burden of upper respiratory illnesses among college and university students: 2002–2003 and 2003–2004 cohorts. Vaccine. 2006 Nov 10;24(44-46):6724-5.

10. Nichol KL, D'Heilly S, Ehlinger EP. Influenza vaccination among college and university students: impact on influenzalike illness, health care use, and impaired school performance. Archives of pediatrics & adolescent medicine. 2008 Dec 1;162(12):1113-8.

11. Iuliano AD, Reed C, Guh A, Desai M, Dee DL, Kutty P, Gould LH, Sotir M, Grant G, Lynch M, Mitchell T. Notes from the field: outbreak of 2009 pandemic influenza A (H1N1) virus at a large public university in Delaware, April-May 2009. Clinical infectious diseases. 2009 Dec 15;49(12):1811-20

12. Nichol KL, Tummers K, Hoyer-Leitzel A, Marsh J, Moynihan M, McKelvey S. Modeling seasonal influenza outbreak in a closed college campus: impact of pre-season vaccination, in-season vaccination and holidays/breaks. PloS one. 2010 Mar 4;5(3):e9548.

13. Guh A, Reed C, Gould LH, Kutty P, Iuliano D, Mitchell T, Dee D, Desai M, Siebold J, Silverman P, Massoudi M. Transmission of 2009 pandemic influenza A (H1N1) at a public university—Delaware, April–May 2009. Clinical infectious diseases. 2011 Jan 1;52(suppl_1):S131-7.

14. Fung IC, Cairncross S. How often do you wash your hands? A review of studies of hand-washing practices in the community during and after the SARS outbreak in 2003. International journal of environmental health research. 2007 Jun 1;17(3):161-83.

15. Keller JJ, Kim JH, Lau JC, Wong AH, Griffiths SM. Intention to engage in preventive behaviors in response to the A/H1N1 pandemic among university entrants in four Chinese cities. Asia Pacific Journal of Public Health. 2014 Jan;26(1):42-7.

16. Perez V, Uddin M, Galea S, Monto AS, Aiello AE. Stress, adherence to preventive measures for reducing influenza transmission and influenza-like illness. J Epidemiol Community Health. 2012 Jul 1;66(7):605-10.
17. Monk-Turner E, Edwards D, Broadstone J, Hummel R, Lewis S, Wilson D. Another look at hand-washing behavior. Social Behavior and Personality: an international journal. 2005 Jan 1;33(7):629-34.
18. Thumma J, Aiello AE, Foxman B. The association between handwashing practices and illness symptoms among college students living in a university dormitory. American journal of infection control. 2009 Feb 1;37(1):70-2.
19. Universities UK. Actions being taken by Universities in response to coronavirus [Internet]. [cited 10 August 2020]. Available from: https://www.universitiesuk.ac.uk/news/Pages/Actions-being-taken-by-universities-in-response-to-coronavirus-.aspx
20. Fernandez AA, Shaw GP. Academic Leadership in a Time of Crisis: The Coronavirus and COVID-19. Journal of Leadership Studies. 2020 May;14(1):39-45.
21. Park D, Kim WG, Choi S. Application of social media analytics in tourism crisis communication. Current Issues in Tourism. 2019 Sep 14;22(15):1810-24.
22. McIntyre JJ, Spence PR, Lachlan KA. Media use and gender differences in negative psychological responses to a shooting on a university campus. Journal of School Violence. 2011 Jul 1;10(3):299-313.
23. Roshan M, Warren M, Carr R. Understanding the use of social media by organisations for crisis communication. Computers in Human Behavior. 2016 Oct 1;63:350-61.

24. Coombs WT. Protecting organisation reputations during a crisis: The development and application of situational crisis communication theory. Corporate reputation review. 2007 Sep 1;10(3):163-76.
25. Veil SR, Buehner T, Palenchar MJ. A work-in-process literature review: Incorporating social media in risk and crisis communication. Journal of contingencies and crisis management. 2011 Jun;19(2):110-22.
26. Triantafillidou A, Yannas P. Social media crisis communication in racially charged crises: Exploring the effects of social media and image restoration strategies. Computers in Human Behavior. 2020 May 1;106:106269.
27. Rutter R, Roper S, Lettice F. Social media interaction, the university brand and recruitment performance. Journal of Business Research. 2016 Aug 1;69(8):3096-104.
28. Pringle J, Fritz S. The university brand and social media: using data analytics to assess brand authenticity. Journal of Marketing for Higher Education. 2019 Jan 2;29(1):19-44.
29. Peruta A, Shields AB. Social media in higher education: Understanding how colleges and universities use Facebook. Journal of Marketing for Higher Education. 2017 Jan 2;27(1):131-43.
30. Linvill DL, McGee SE, Hicks LK. Colleges' and universities' use of Twitter: A content analysis. Public Relations Review. 2012 Nov 1;38(4):636-8.
31. Lund B. Universities engaging social media users: an investigation of quantitative relationships between universities' Facebook followers/interactions and university attributes. Journal of Marketing for Higher Education. 2019 Jul 3;29(2):251-67.

32. WONKHE. How should providers be communicating about Covid-19? [Internet], 2020. Available from: https://wonkhe.com/blogs/how-should-providers-be-communicating-about-covid-19/

33. Hagar C. Using research to aid the design of a crisis information management course. InALISE Annual Conference SIG Multicultural, Ethnic & Humanistic Concerns (MEH). Information Seeking and Service Delivery for Communities in Disaster/Crisis, San Antonio 2006.

34. Hagar C. Crisis Informatics. In Khosrow-Pour M. ed. Encyclopaedia of Information Science and technology, 3rd ed. IGI Global; 2014. p. 1350-1358.

35. Palen L, Vieweg S, Sutton J, Liu SB, Hughes A. Crisis informatics: Studying crisis in a networked world. InProceedings of the Third International Conference on E-Social Science 2007 Oct 7 (pp. 7-9).

36. Palen L, Vieweg S, Liu SB, Hughes AL. Crisis in a networked world: Features of computer-mediated communication in the April 16, 2007, Virginia Tech event. Social Science Computer Review. 2009 Nov;27(4):467-80.

37. Palen L, Hughes AL. Social media in disaster communication. In Handbook of disaster research 2018 (pp. 497-518). Springer, Cham.

38. Gründer-Fahrer S, Schlaf A, Wiedemann G, Heyer G. Topics and topical phases in German social media communication during a disaster. Natural language engineering. 2018 Mar;24(2):221-64.

39. Palen L, Anderson J, Bica M, Castillos C, Crowley J, Díaz P, Finn M, Grace R, Hughes A, Imran M, Kogan M. Crisis Informatics: Human-Centered Research on Tech & Crises 2020.

40. Palen L, Liu SB. Citizen communications in crisis: anticipating a future of ICT-supported public participation. In Proceedings of the SIGCHI conference on Human factors in computing systems 2007 Apr 29 (pp. 727-736).

41. Caragea C, McNeese NJ, Jaiswal AR, Traylor G, Kim HW, Mitra P, Wu D, Tapia AH, Giles CL, Jansen BJ, Yen J. Classifying text messages for the Haiti earthquake. InISCRAM 2011 May 8.

42. Imran M, Castillo C, Diaz F, Vieweg S. Processing social media messages in mass emergency: A survey. ACM Computing Surveys (CSUR). 2015 Jun 26;47(4):1-38.

43. Marinoni G, Van’t Land H, Jensen T. The impact of Covid-19 on higher education around the world. IAU Global Survey Report. 2020 May.

44. Butler-Henderson K, Crawford J, Rudolph J, Lalani K, Sabu KM. COVID-19 in Higher Education Literature Database (CHELD V1): An open access systematic literature review database with coding rules. Journal of Applied Learning & Teaching. 2020;3(2):1-6.

45. Aristovnik A, Keržič D, Ravšelj D, Tomaževič N, Umek L. Impacts of the COVID-19 pandemic on life of higher education students: A global perspective. Sustainability. 2020 Jan;12(20):8438.

46. Crawford J, Butler-Henderson K, Rudolph J, Malkawi B, Glowatz M, Burton R, Magni P, Lam S. COVID-19: 20 countries' higher education intra-period digital pedagogy responses. Journal of Applied Learning & Teaching. 2020;3(1):1-20.

47. Gonzalez T, De La Rubia MA, Hincz KP, Comas-Lopez M, Subirats L, Fort S, Sacha GM. Influence of COVID-19 confinement on students’ performance in higher education. PloS one. 2020 Oct 9;15(10):e0239490.

48. Händel M, Stephan M, Gläser-Zikuda M, Kopp B, Bedenlier S, Ziegler A. Digital readiness and its effects on higher education students' socio-emotional perceptions in the context of the COVID-19 pandemic. Journal of Research on Technology in Education. 2020 Nov 9:1-3.
49. Ye S, Hartmann RW, Söderström M, Amin MA, Skillinghaug B, Schembri LS, Odell LR. Turning Information Dissipation into Dissemination: Instagram as a Communication Enhancing Tool during the COVID-19 Pandemic and Beyond. Journal of Chemical Education. 2020 Aug 17;97(9):3217-22.
50. Bao W. COVID-19 and online teaching in higher education: A case study of Peking University. Human Behavior and Emerging Technologies. 2020 Apr;2(2):113-5.
51. Ayman U, Kaya AK, Kuruç ÜK. The Impact of Digital Communication and PR Models on the Sustainability of Higher Education during Crises. Sustainability. 2020 Jan;12(20):8295.
52. Sobaih AE, Hasanein AM, Abu Elnasr AE. Responses to COVID-19 in higher education: Social media usage for sustaining formal academic communication in developing countries. Sustainability. 2020 Jan;12(16):6520.
53. Cevasco KE, North HM, Zeitoun SA, Wofford RN, Matulis GA, Gregory AF, Hassan MH, Abdo AD, Farris D, Roess AA, von Fricken ME. COVID-19 observations and accompanying dataset of non-pharmaceutical interventions across US universities, March 2020. PloS one. 2020 Oct 16;15(10):e0240786.
54. Viner RM, Russell SJ, Croker H, Packer J, Ward J, Stansfield C, Mytton O, Bonell C, Booy R. School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. The Lancet Child & Adolescent Health. 2020 Apr 6.

55. Gressman PT, Peck JR. Simulating COVID-19 in a university environment. Mathematical Biosciences. 2020 Oct; 328
56. Lopman B, Liu C, Le Guillou A, Handel A, Lash TL, Isakov A, Jenness S. A model of COVID-19 transmission and control on university campuses. medRxiv. 2020 Jan 1.
57. Knight M. Pandemic Communication: A New Challenge for Higher Education. Business and Professional Communication Quarterly. 2020 May;83(2):131-132.
58. Mackert M, Table B, Yang J, Bouchacourt L, Woods JM, Bernhardt JM, Wagner JH. Applying Best Practices from Health Communication to Support a University's Response to COVID-19. Health communication. 2020 Dec 5;35(14):1750-3.
59. Duong V, Pham P, Yang T, Wang Y, Luo J. The ivory tower lost: How college students respond differently than the general public to the covid-19 pandemic. arXiv preprint arXiv:2004.09968. 2020 Apr 21.
60. Mossa-Basha M, Medverd J, Linnau K, Lynch JB, Wener MH, Kicska G, Staiger T, Sahani D. Policies and guidelines for COVID-19 preparedness: experiences from the University of Washington. Radiology. 2020 Apr 8:201326.
61. Charoensukmongkol P, Phungsoonthorn T. The Interaction Effect of Crisis Communication and Social Support on The Emotional Exhaustion of University Employees during the COVID-19 Crisis. International Journal of Business Communication. 2020 Jan 1:2329488420953188.
62. Charoensukmongkol P, Phungsoonthorn T. The effectiveness of supervisor support in lessening perceived uncertainties and emotional exhaustion of university employees during the COVID-19 crisis: the constraining role of organisational intransigence. The Journal of general psychology. 2020 Jul 20:1-20.

63. Sergent K, Stajkovic AD. Women's leadership is associated with fewer deaths during the COVID-19 crisis: Quantitative and qualitative analyses of United States governors. Journal of Applied Psychology. 2020 Jul 2.

64. Bootsma MC, Ferguson NM. The effect of public health measures on the 1918 influenza pandemic in US cities. Proceedings of the National Academy of Sciences. 2007 May 1;104(18):7588-93.

65. Coombs WT. Protecting organisation reputations during a crisis: The development and application of situational crisis communication theory. Corporate reputation review. 2007 Sep 1;10(3):163-76.

66. Coombs WT. The value of communication during a crisis: Insights from strategic communication research. Business Horizons. 2015 Mar 1;58(2):141-8.

67. Huang Y, DiStaso M. Responding to a Health Crisis on Facebook: The Effects of Response Timing and Message Appeal. Public Relations Review. 2020 May 14:101909.

68. Claeys AS, Cauberghe V. Crisis response and crisis timing strategies, two sides of the same coin. Public Relations Review. 2012 Mar 1;38(1):83-8.

69. Arpan LM, Roskos-Ewoldsen DR. Stealing thunder: Analysis of the effects of proactive disclosure of crisis information. Public Relations Review. 2005 Sep 1;31(3):425-33.

70. Boehmke FJ, Witmer R. Disentangling diffusion: The effects of social learning and economic competition on state policy innovation and expansion. Political Research Quarterly. 2004 Mar;57(1):39-51.

71. Neumayer E, Plümper T. Robustness tests for quantitative research. Cambridge University Press; 2017 Aug 17.

72. Gov UK. Check if a university or college is officially recognized [Internet]. [cited 9 April 2020]. Available from: https://www.gov.uk/check-a-university-is-officially-recognised/recognised-bodies

73. Sauerbrei W, Royston P, Bojar H, Schmoor C, Schumacher M. Modelling the effects of standard prognostic factors in node-positive breast cancer. British Journal of Cancer. 1999 Apr;79(11):1752-60.

74. Drayton E, Waltmann B. Will universities need a bailout to survive the COVID-19 crisis? [Internet]. July 2020 [cited 7 July 2020]. Available from: https://www.ifs.org.uk/uploads/BN300-Will-universities-need-bailout-survive-COVID-19-crisis.pdf

75. Finkelstein S. Power in top management teams: Dimensions, measurement, and validation. Academy of Management journal. 1992 Aug 1;35(3):505-38.

76. Hambrick DC, Mason PA. Upper echelons: The organisation as a reflection of its top managers. Academy of management review. 1984 Apr 1;9(2):193-206.

77. Universities UK. Our members [Internet]. [cited 28 October 2020]. Available from: https://www.universitiesuk.ac.uk/about/Pages/member-institutions.aspx

78. Higher Education Statistics Authority. Open data and official statistics [Internet]. [cited 30 April 2020]. Available from: https://www.hesa.ac.uk/data-and-analysis

79. Park S, Hendry DJ. Reassessing Schoenfeld residual tests of proportional hazards in political science event history analyses. American Journal of Political Science. 2015 Oct;59(4):1072-87.

80. Jin S, Boehmke FJ. Proper Specification of Nonproportional Hazards Corrections in Duration Models. Political Analysis. 2017 Jan;25(1):138-44.

81. Berry FS, Berry WD. State lottery adoptions as policy innovations: An event history analysis. The American Political Science Review. 1990 Jun 1:395-415.
82. Mooney CZ. Modeling regional effects on state policy diffusion. Political Research Quarterly. 2001 Mar;54(1):103-24.
83. Gilardi F, Füglister K. Empirical modeling of policy diffusion in federal states: the dyadic approach. Swiss Political Science Review. 2008 Sep;14(3):413-50.
84. Boehmke FJ. Policy emulation or policy convergence? Potential ambiguities in the dyadic event history approach to state policy emulation. The Journal of Politics. 2009a Jul;71(3):1125-40.
85. Boehmke FJ. Approaches to modeling the adoption and diffusion of policies with multiple components. State Politics & Policy Quarterly. 2009b Jun;9(2):229-52.
86. Shipan CR, Volden C. Policy diffusion: Seven lessons for scholars and practitioners. Public Administration Review. 2012 Nov;72(6):788-96.
87. Graham ER, Shipan CR, Volden C. The diffusion of policy diffusion research in political science. British Journal of Political Science. 2013 Jul 1:673-701.
88. Adolph C, Amano K, Bang-Jensen B, Fullman N, Wilkerson J. Pandemic politics: Timing state-level social distancing responses to COVID-19. medRxiv. 2020 Jan 1.
89. Volden C. States as policy laboratories: Emulating success in the children's health insurance program. American Journal of Political Science. 2006 Apr;50(2):294-312.
90. Padgham M, Sumner DM, Karney CFF. Geodist: Fast, Dependency-Free Geodesic Distance Calculations [Internet]. [cited 16 May 2020]. Available from: https://cran.case.edu/web/packages/geodist/index.html
91. Gov UK. Coronavirus (COVID-19) in the UK [Internet]. [cited 1 May 2020]. Available from: https://coronavirus.data.gov.uk/about-data#legacy-csv-downloads

92. Tom EW. Covid-19 UK data [Internet] cited 10 August 2020, Available from: https://github.com/tomwhite/covid-19-uk-data

93. Office of National Statistics. Counties and Unitary Authorities (December 2017) Full Clipped Boundaries in UK [Internet]. [cited 10 May 2020]. Available from: https://geoportal.statistics.gov.uk/datasets/6638c31a8e9842f98a037748f72258ed_0

94. Pebesma E, Bivand R. Classes and methods for spatial data in R [Internet]. [cited 16 May 2020]. Available from: https://cran.r-project.org/web/packages/sp/index.htmll

95. Beck N, Katz JN, Tucker R. Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. American Journal of Political Science. 1998 Oct 1;42(4):1260-88.

96. Carter DB, Signorino CS. Back to the future: Modeling time dependence in binary data. Political Analysis. 2010;18(3):271-92.

97. Keefer P, Neumayer E, Plümper T. Earthquake propensity and the politics of mortality prevention. World Development. 2011 Sep 1;39(9):1530-41.

98. Quiroz Flores A, Smith A. Leader survival and natural disasters. British Journal of Political Science. 2013 Oct 1:821-43.

99. Neumayer E, Plümper T, Barthel F. The political economy of natural disaster damage. Global Environmental Change. 2014 Jan 1;24:8-19.

100. Plümper T, Flores AQ, Neumayer E. The double-edged sword of learning from disasters: mortality in the Tohoku tsunami. Global Environmental Change. 2017 May 1;44:49-56.

101. OBR. Coronavirus reference scenario: The OBR's coronavirus analysis [Internet]. 2020. Available from: https://obr.uk/coronavirus-analysis/

Supporting information

S1 Table. Verified Twitter accounts.

S2 Table. Additional controls for university leadership for Model 1 in Table 1 and Model 1 in Table 3.

S3 Table. Summary statistics.

S4 Appendix. Specific HESA tables.

S5 Table. Additional lags for Model 2 in Table 2 (dyadic unconditional).

S6 Table. Alternative definitions of a neighbourhood for dyadic models of Table 2.

S7 Table. Additional lags for Model 2 in Table 4 (dyadic unconditional).

S8 Table. Alternative definitions of a neighbourhood for dyadic models of Table 4.