Supporting Information

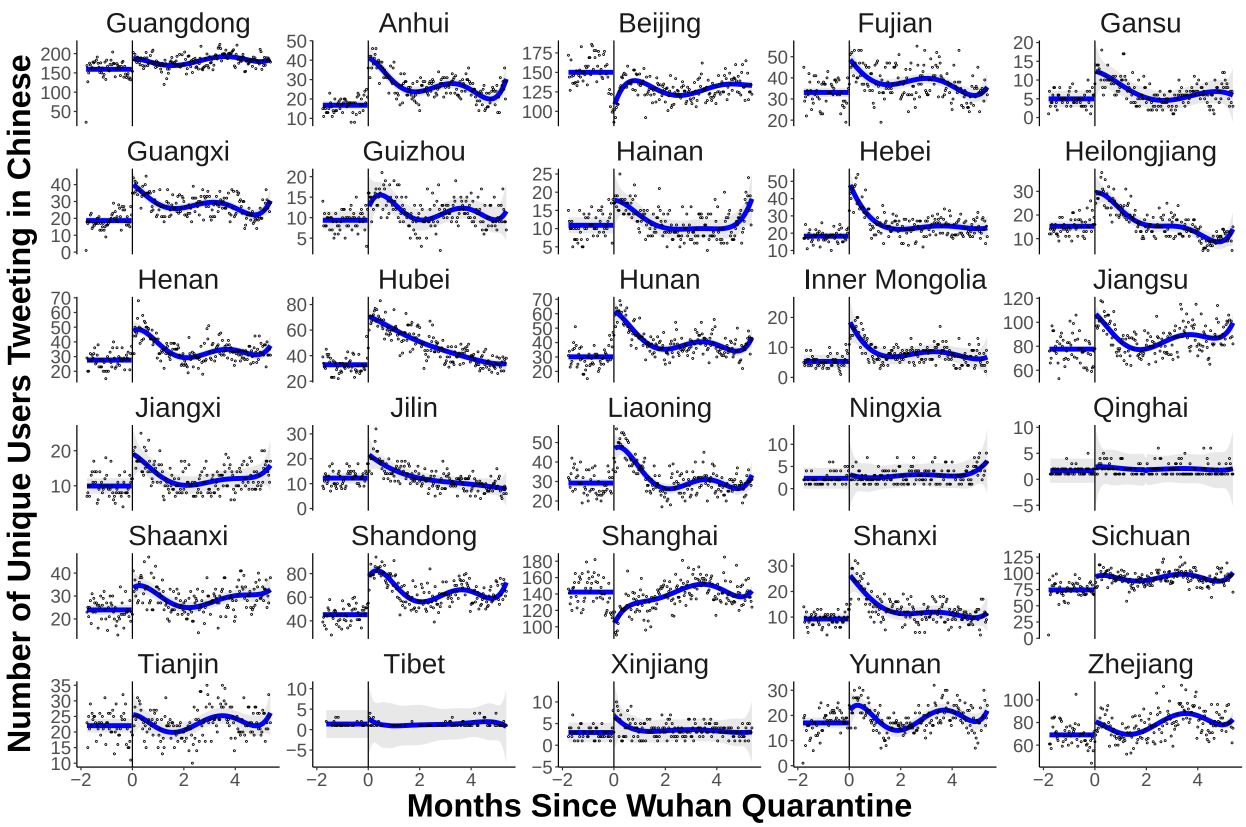
COVID-19 Increased Censorship Circumvention And Access To Sensitive Topics In China

Contents

1 Twitter Activity by Province

Figure A1 shows the number of unique, geolocating users who are tweeting in Chinese by province. The x-axis is the number of months before (negative) or after (positive) the initial coronavirus lockdown in Hubei province. The blue line is a pre-lockdown average for x less than 0 and a five term polynomial regression for x greater than or equal to 0 (where 0 is the first day of Hubei's lockdown). The points in Figure 2 are the values of the blue line by province for x equals 1/30 (first day of lockdown) and x equals 1 (day 30 of lockdown).

Figure A1: Increases in Geolocated Twitter Activity by Province (modeled)



2 Twitter Data

From a global sample of tweets with GPS coordinates, we found the 1,448,850 tweets from China from December 1, 2019 through June 30, 2020, 367,875 of which are in Chinese. This corpus contains 101,553 unique users, 43,114 of whom had names or descriptions in Chinese. These dates were chosen to encompass a baseline period and the height of COVID-19 in China. This corpus is used for Figures 2 and 4, evaluating the impact of the lockdown on tweeting behavior.

For the follower analysis (Figures 5 and 6), we sample 5,000 of these 43,114 accounts. For these 5,000 random users in China, we download who they follow, their "friends" in Twitter parlance. From these friends, we identify the 5,000 most commonly followed accounts that are either a Chinese language account or have Chinese characters in their name or description field. Of these 5,000 most common friends, the vast majority were pornography accounts. We therefore hand-categorized the accounts into pornography or not pornography. We keep the 354 non-porn accounts and sample 200 from the remaining 4,646 porn accounts.

We then download the followers of these 554 accounts. We identify 38,050,454 total followers. For each, we identify the location of the users. Because very few of these followers have geolocated information, we rely on the language of their Twitter status and their self-reported location to distinguish between mainland and overseas followers. We only include users whose status language is Chinese in order to study only Chinese language followers of these accounts. Followers are classified as Mainland Chinese if the location field contains the name of a Chinese city, town, or province. Followers are classified as from Hong Kong if the location field contains the name of a district in Hong Kong. Followers are classified as Taiwanese if the location field contains the name of a Taiwanese city, county, or district. Followers are classified as US if if the location field contains the name of states or state abbreviations (in capital letters).

3 Mobility and Twitter Usage

To better understand the relationship between lockdown and Twitter usage, we use the publicly-available human mobility data from Baidu Qianxi (`https://qianxi.baidu.com/2020/`), which tracks real-time migration (including moves in & out of provinces and within city movements) across China during the Lunar New Year period in both 2020 and 2019. The move out data is downloaded from Harvard Dataverse (China Data Lab 2020; Hu et al. 2020), and we scrape the within-city movement data from Baidu Qianxi.

Figure A2 plots the average within city movement index in both 2020 (real black line) and 2019 during the same period in the Chinese Lunar New Year (dotted line). Specifically, since the New Years day is on February 5 in 2019 and January 25 in 2020, we shifted the dates in 2019 backwards for 12 days to match the dates in 2020. Red vertical line indicates the day of Wuhan lockdown. One can see that almost all provinces experienced a huge decrease in human mobility after January 23 in 2020, compared to the same period in 2019. In 2019, we only see significant decreases in mobility in Beijing, Shanghai, and Tianjin.

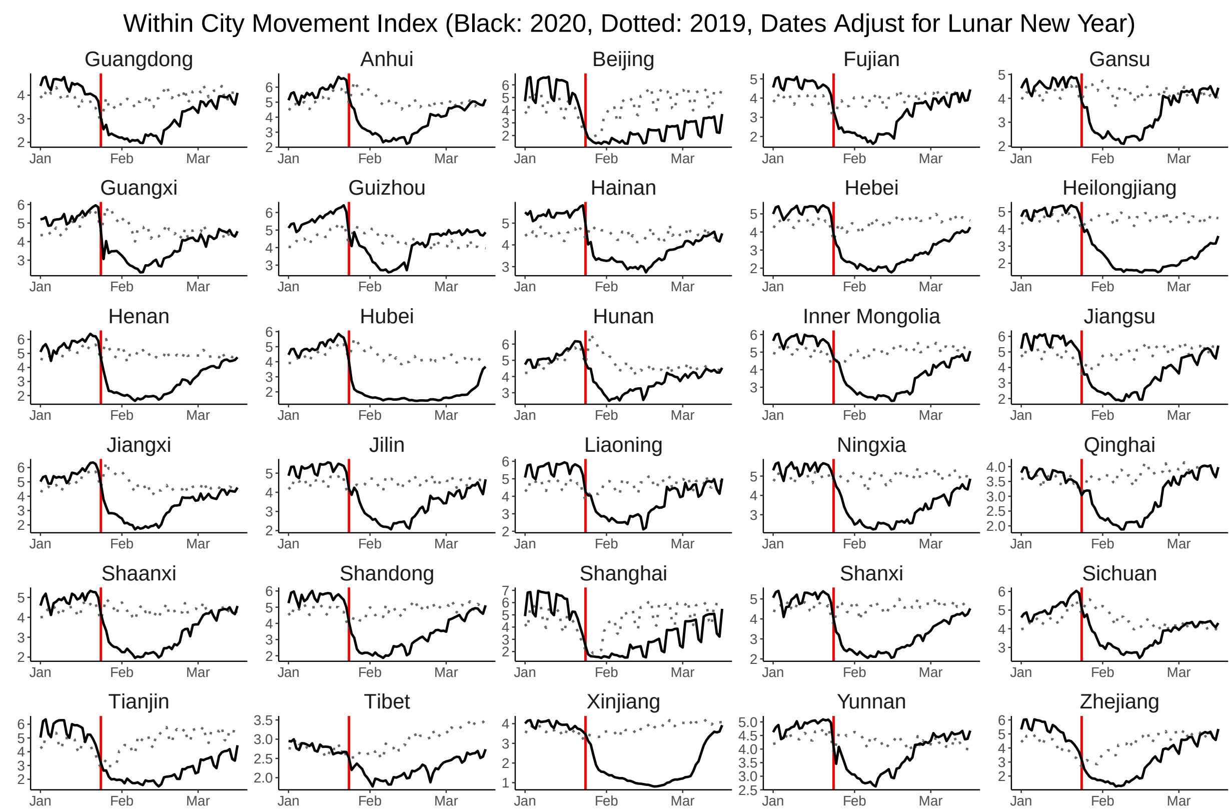


Figure A2: Within city movement index by Province (black: 2020, dotted: same period in 2019).
Note: Real black line indicates the time series for the average within city movement index by province in 2020. Dotted line indicates the average within city movement index by province during the same Chinese Lunar New Year period in 2019. Red vertical line indicates the day of Wuhan lockdown. Chongqing is excluded since it is not counted in Twitter's geolocation map.

We also validate that the increase in geolocated Twitter users is correlated with the decrease in human mobility. The left panel of Figure A3 plots this correlation. Let $M_{i,t}$ denote the mobility

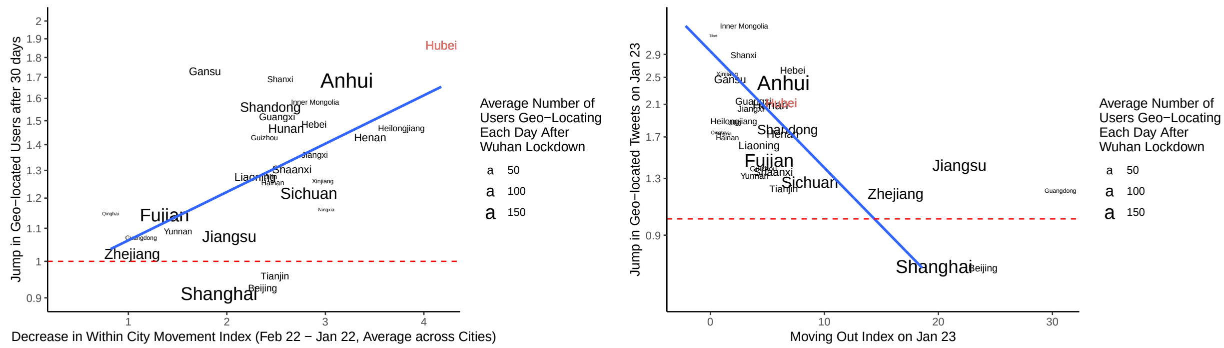


Figure A3: Reduction in within city movement and increase in geolocated Twitter users during the month of Wuhan lockdown (left); degree of moving out and increase in geolocated Twitter users on the day of Wuhan lockdown (right).

Note: The left panel plots the correlation between decreased mobility and increased geolocated Twitter users during the first 30 days of Wuhan lockdown. The x-axis plots the decrease in within city movement index from January 22 (the day before Wuhan lockdown) to Feb 22. The y-axis plots the increase in geolocated Twitter users for 30 days after the Wuhan lockdown compared to the average number of geolocated Twitter users in a province before the Wuhan lockdown. The right panel plots the relationship between moving out of province and the increase of geolocated Twitter users on January 23, the day of Wuhan lockdown. Estimates and day of lockdown are drawn from a five term polynomial regression on the number of unique geolocated Twitter users per day after the lockdown. These province-by-province polynomials are displayed over the raw data in Figure A1.

index for province i on date t. The x-axis plots the decrease in within city movement index from January 22 (the day before the Wuhan lockdown) to February 22, $M_{i,\text{Jan22}} - M_{i,\text{Feb22}}$. The y-axis plots the increase in geolocated Twitter users for 30 days after the Wuhan lockdown compared to the average number of geolocated Twitter users in a province before the Wuhan lockdown. This shows that the more reduction in human mobility, the more increase in geolocated Twitter users, comparing to the levels before the lockdown. Hubei province experience the most reduction in mobility and most increase in the number of geolocated Twitter users.

In Figure A1 we see increases in geolocated Twitter users in most provinces except Beijing and Shanghai. One explanation for this is that Twitter users in Beijing and Shanghai left the cities during the outbreak. Mobility data supports this explanation. The right panel of Figure A3 plots the relationship between moving out of a province on January 23, the day of Wuhan lockdown, and the increase of number of geolocated Twitter users on the same day, compared to the average number of geolocated Twitter users in a province before the lockdown. One can see that the more people moving out, the less jump in Twitter user on the day of lockdown. Beijing, Shanghai, and Guangdong all experienced large outflows of individuals on the day of Wuhan lockdown.

Since the period of Wuhan lockdown overlaps with the Chinese Lunar New Year, increased Twitter usage could partly be due to general boredom during the New Year. To explore New Year versus pandemic effects, we normalize both mobility and number of Twitter users in 2020 by those in the same period in 2019. To do so, we first adjust the dates in 2019 backwards for 6 days to match the dates of 2020 Lunar New Year. Then, we create normalized mobility and Twitter usage. Specifically, denote $M_{i,y,t}$ the mobility index and $T_{i,y,t}$ the Twitter usage for province i in year y on date t. The normalized mobility index would be

$$M_{i,2020,t}/M_{i,2019,t}$$

and the normalized Twitter usage would be

$$T_{i,2020,t}/T_{i,2019,t}.$$

We then plot the weekly change in mobility and Twitter usage after Wuhan lockdown, comparing to the period before Wuhan lockdown. Figure A4 shows the plots. In mathematical notations, for the first week of Wuhan lockdown, we plot

$$\frac{M_{i,2020,\text{Week 1}}/M_{i,2019,\text{Week 1}}}{M_{i,2020,\text{Week 0}}/M_{i,2019,\text{Week 0}}}$$

on the x-axis and

$$\frac{T_{i,2020,\text{Week 1}}/T_{i,2019,\text{Week 1}}}{T_{i,2020,\text{Week 0}}/T_{i,2019,\text{Week 0}}}$$

on the y-axis. This is shown in the top left panel in Figure A4. The other panels shows the corresponding ratios for the 2nd, 3rd, and 4th week, respectively (all relative to the week before lockdown).

Figure A4 shows that we still find effects of reduced mobility on increased Twitter usage, after adjusting for the decrease in movement driven purely from New Year, in the early periods of lockdown (at least for the first week of lockdown, the correlation for the second week is not statistically significant). In the 3rd and 4th weeks, we find a general increase in Twitter usage in most provinces, regardless of the relative decrease in mobility in these weeks. In other words, the mobility-induced effect specific to Wuhan lockdown fades out in around 2 weeks, and there's a general increase in Twitter usage across China that is not related to reduced human mobility. This pattern suggests that the increase in Twitter usage is not driven only by people's staying at home because, if that is the case, we expect to see a continued relationship between relative reduction in mobility and increase in Twitter usage, as other Provinces started to announce stay-at-home orders. This pattern is also not driven only by New Year because we should not expect to see an overall increase in Twitter usage after normalizing with the same New Year period in 2019.

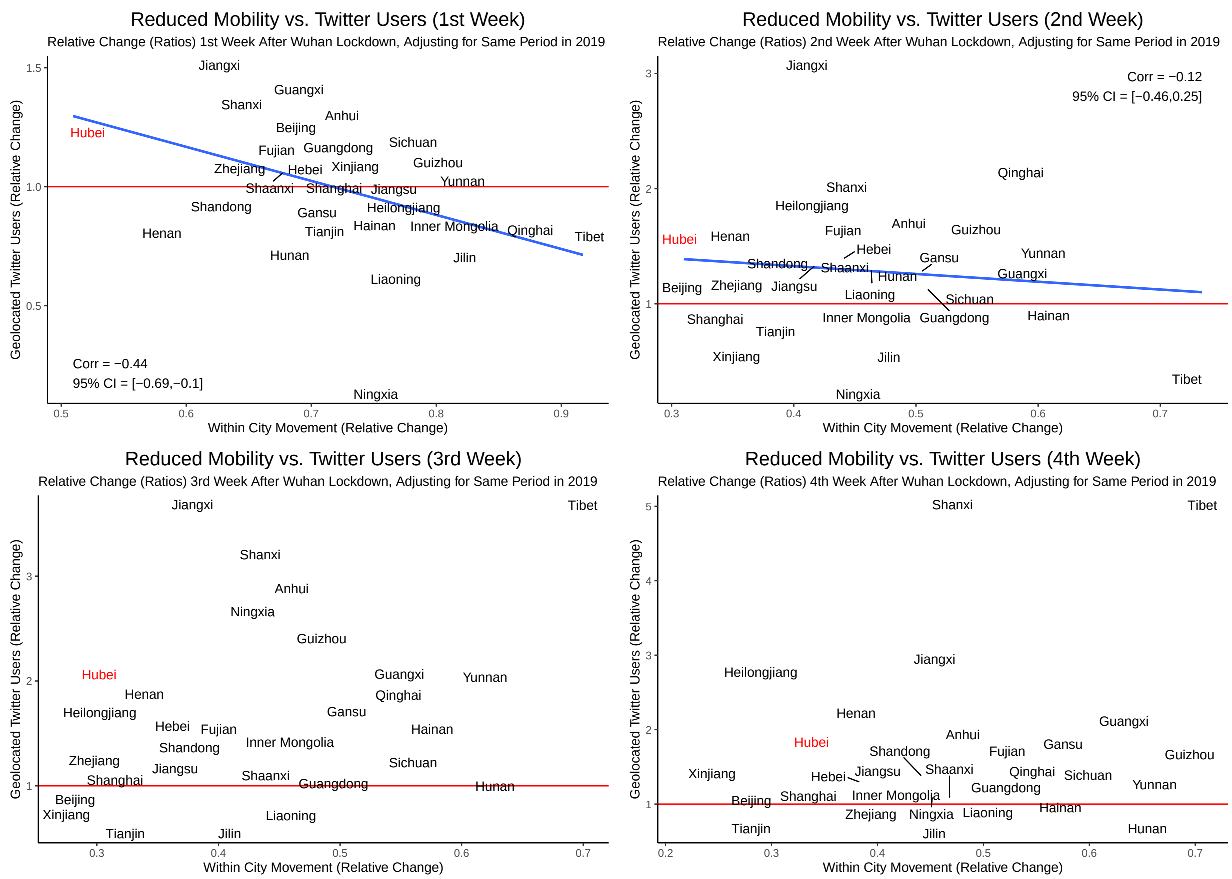
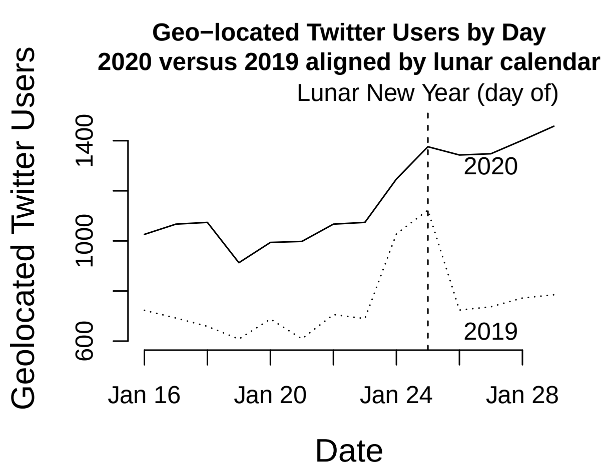


Figure A4: Weekly changes in within city movement and geolocated Twitter users relative to pre-lockdown period, after adjusting for the same period in 2019.

Note: We plot the weekly relative change in mobility and Twitter usage after Wuhan lockdown, adjusting for the same period in 2019 and comparing to the period before Wuhan lockdown. Specifically, denote $M_{i,y,t}$ the mobility index and $T_{i,y,t}$ the Twitter usage for province i in year y on date t. For the first week of Wuhan lockdown, we plot (in the top left panel) $\frac{M_{i,2020,\text{Week }1}/M_{i,2019,\text{Week }1}}{M_{i,2020,\text{Week }0}/M_{i,2019,\text{Week }0}}$ on the x-axis and $\frac{T_{i,2020,\text{Week }1}/T_{i,2019,\text{Week }1}}{T_{i,2020,\text{Week }0}/T_{i,2019,\text{Week }0}}$ on the y-axis. The other panels shows the same for the 2nd, 3rd, and 4th weeks, respectively.



Note also that the first week of Twitter use in 2020 was not much higher than in 2019 because 2019 saw a very large number of posts on Chinese Lunar New Year. This increase was presumably related to New Year related posts, and these celebratory posts did not increase to the same extent during the start of the COVID-19 pandemic.

4 Effect Size

4.1 New Twitter Users

This section provides rough estimates of absolute increases in Twitter use in China, and sections below expand it to consider increased Twitter followings and increased Wikipedia use. Note that these are estimates for increased usage on only these sites, which require that new users from Mainland China (where these sites are blocked) 1) create an account to view Twitter content and 2) use cookies (to be recorded in the Wikipedia unique device data). Other sites that do not require accounts could have seen larger increases, and the Wikipedia unique device counts are underestimates.

The top panel of Figure 2 shows a 10% long-term increase in the number of geolocating users from China. In 2019, as reported in (Mozur 2019), Professor Daniela Stockman of the Hertie School of Governance surveyed 1,627 internet users in China and found .4% of them use Twitter; the article reports that number as 3,200,000. Roughly, if the same 10% increase applies to all users from China and the long-term increase reflects a new pool of users (the number of unique geotagging users in our sample in May 2020 was around 10% higher than in December 2019), then 320,000 new users joined Twitter because of the crisis.

We can assess this estimate by considering 1) the fraction of (posting) users who geotag and 2) the number of unique geotagging users in our data. For 1), using a sample of 100 hours of non-geolocated tweets from 2019.01.01-2020.12.31, we found 37,957 in Chinese. Assigning location to these tweets using the same code that was used to assign location to followers of the most commonly followed accounts, we then found that 1.79% of tweets and 1.95% of users from China geotag. For 2), we find that 47,389 unique Twitter users geotagged (in Chinese and in China) in our sample (note, however, that our 1% sample captures approximately 56% of tweets that are geotagged). Dividing this number by 0.0195 gives us 2.4 million Twitter users, suggesting that somewhere around 70% ($\frac{2.4}{3.2\times 1.1}$) of geotagging Twitter users in China publicly geotagged posts and were in our sample.

Though this number is small in the context of China's 1.4 billion inhabitants, it is nonetheless important for three reasons. First, the effects in this paper are a minimum effect size for Twitter since accounts do not have to use geolocation or provide an accurate self-reported location in their profile. Second, the effects documented herein focus only on one banned platform (Twitter) and website (Wikipedia), and there is no reason to think the same behavior did not occur on other banned platforms like Facebook, Telegram, and Instagram as well as banned websites such as Reddit or *The New York Times*. Third, the Chinese government behaves as though these relatively small numbers threaten it. Since 2018, it has become increasingly repressive in response to comments its citizens make on platforms unavailable in China. Recently, several individuals from China have been arrested for comments made on platforms such as WhatsApp (owned by Facebook, unavailable inside the Great Firewall) and Twitter (Mozur 2018). The government has also started large influence campaigns on social media platforms that are unavailable domestically, including Twitter (Kinetz 2021). If the behaviors documented in this paper were immaterial, then we believe the government would not put such a priority on attempting to control speech on these platforms.

4.2 Followers

In addition to causing new people to join Twitter, the crisis caused more people to follow accounts posting sensitive content. Here, we estimate the number of surplus followers from China and show that they persist after the crisis, perhaps at greater rates than users who follow after.

Figure A5 shows the absolute number of excess followers (top) and its ratio (bottom). The absolute number is the total number of new followers minus the total number of predicted new followers based on the December daily average growth rate per category; the bottom panel divides the new follower count by the predicted number of new followers. Several interesting patterns emerge. First, the crisis clearly causes all account types to gain followers; some, such as pornography and international news agencies, may even have served as early warning indicators since they receive excess followers before the Wuhan lockdown. Second, the categories with the most excess followers, citizen journalists/political bloggers and international news agencies, are exactly those people would seek out in a crisis. By the end of March, 53,860 more accounts follow citizen journalists/political bloggers than would have happened without the crisis; for international news agencies, 52,144. Third, normalizing for the expected number of new followers reinforces that attention was paid to sensitive categories. Extra, early attention is paid to the citizen journalists and activist categories (which received almost 4 times as many new followers during the lockdown as we would expect based on December's following rate), while international news agencies' importance decreases to third place. Normalizing emphasizes the increased attention activists receive since they have relatively fewer followers than the other categories. Fourth, Chinese accounts increase their following of state media or Chinese officials once Hubei's lockdown lifts, though from a low base.

Importantly, these excess followers persist a year after the lockdown. To make this claim, we crawled the follower list of the same popular accounts starting on May 31, 2021, more than one year after the first crawl, and assigned location using the same procedure as before. Comparing the 2021 follower lists to 2020 shows which followers stopped following the popular accounts. We then calculate the percentage of the 2020 followers that persist in 2021 by account type, follower location, and date. Table A1 shows these results.

Table A1: Persistence of Followers by Account Type and Period Following Starts

	Pre-Lockdown			Lockdown			Post-Lockdown		
	China	Hong Kong	Taiwan	China	Hong Kong	Taiwan	China	Hong Kong	Taiwan
International News Agencies	87.31	87.71	85.51	90.80	90.19	83.20	89.09	87.35	88.70
Citizen Journalists / Political Bloggers	72.84	79.58	78.73	87.49	86.35	85.00	81.69	83.01	79.64
Activists or US / Taiwan / Hong Kong Politics	78.27	76.74	76.45	88.02	86.40	83.83	85.82	85.46	83.99
Pornography Accounts	85.56	84.28	83.00	88.84	87.51	89.52	86.32	87.19	86.88
State Media or Chinese Officials	82.72	81.38	84.62	87.99	86.36	84.94	85.90	86.61	82.45
Non-Political Bloggers or Entertainment Accounts	73.75	72.94	65.66	87.80	87.01	87.19	81.90	85.13	83.61

Note: Each cell is the percent of followers from April 2020 that still follow the six account types (row) in May 2021, by follower location and period the follower started following the account. The lockdown period is January 23, 2020 - March 13, 2020. Post-lockdown refers to March 14-April 1.

Accounts from China that start following the popular accounts during the lockdown period persist at the same to slightly higher rates than those that start following before or after then. 87.31% of accounts from China that start following international news agencies before the lockdown persist versus 89.09% that start following after the lockdown. The difference is especially stark for citizen journalists/political bloggers. Finally, since older followers should have a lower persistence rate

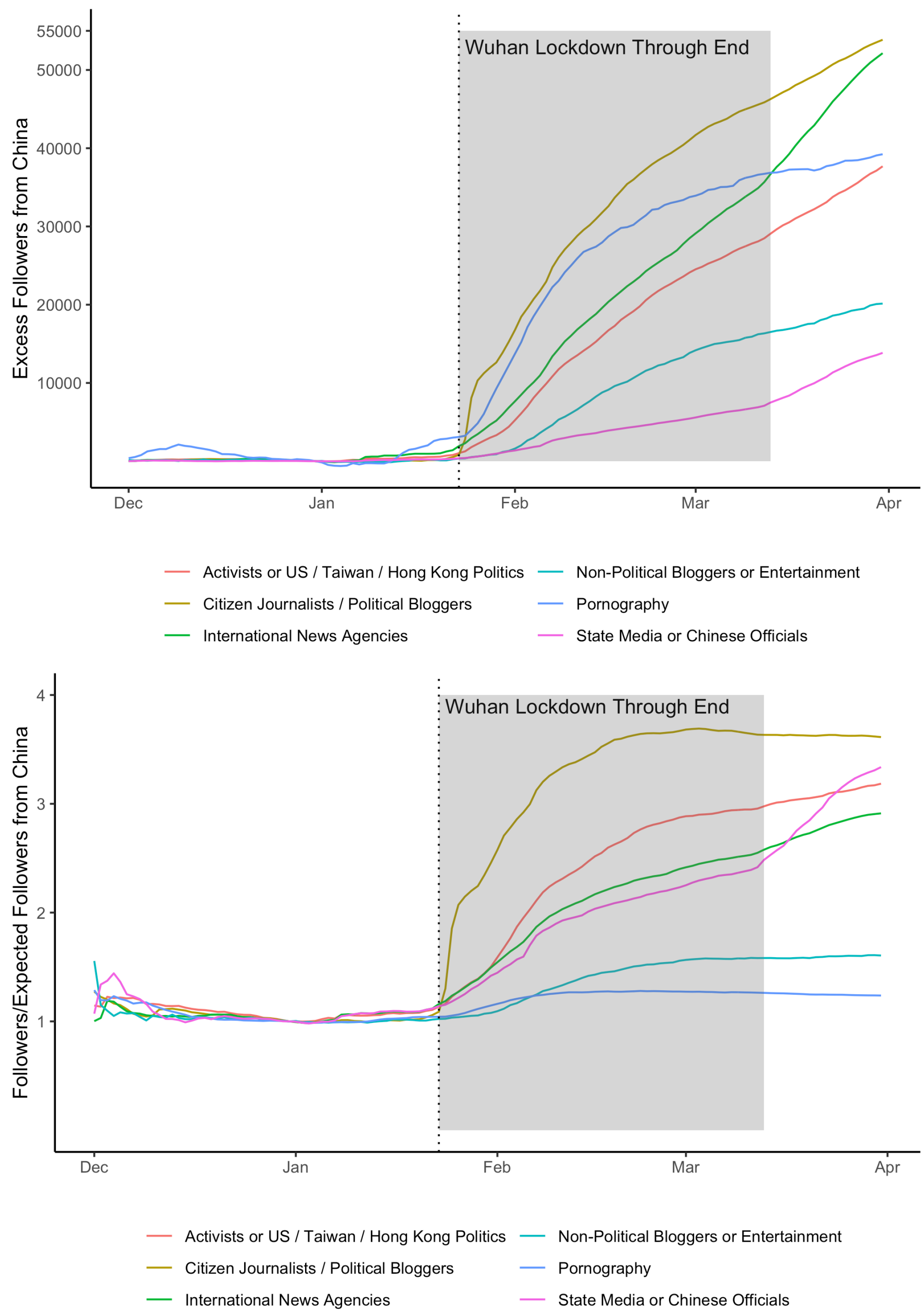


Figure A5: Excess Followers, absolute (top) and ratio normalized by category growth rate (bottom). Growth rate is calculated based on the December 2019 average number of new followers by category.

since more time has passed, it is striking that accounts that start following during lockdown have higher persistence rates than newer accounts, those that start following during the seventeen days after the lockdown ends. The increased exposure to sensitive content persists after the crisis passes at rates equal to or greater than for non-crisis periods.

4.3 Number of unique devices accessing Wikipedia with cookies enabled

Wikipedia tracks the number of unique devices that have accessed its site each day and month using a 'privacy-sensitive access cookie' (`https://dumps.wikimedia.org/other/analytics/`). By design, this number does not count devices not accepting cookies through private browsing (as we might expect from users accessing Wikipedia from within Mainland China) and so underestimates access (see `https://diff.wikimedia.org/2016/03/30/unique-devices-dataset/`). However, this estimate still provides some perspective on the number of individuals who might be accessing the Chinese language version of Wikipedia over time. For the Chinese language version of Wikipedia, 40.8 million devices accessed the site during December 2019 and 42.8 million per month during January, February, and March 2020, an increase of approximately 2 million devices. 3.34 million devices accessed the Chinese language Wikipedia per day in December 2019 and 3.66 million accessed the site per day during lockdown, an increase of approximately 300 thousand devices. These differences are somewhat smaller when comparing to the last half of 2019 (during ongoing protests in Hong Kong) – an increase of 1 million unique devices monthly during lockdown compared to July through December 2019, and an increase of 200 thousand devices daily.

5 Robustness Checks

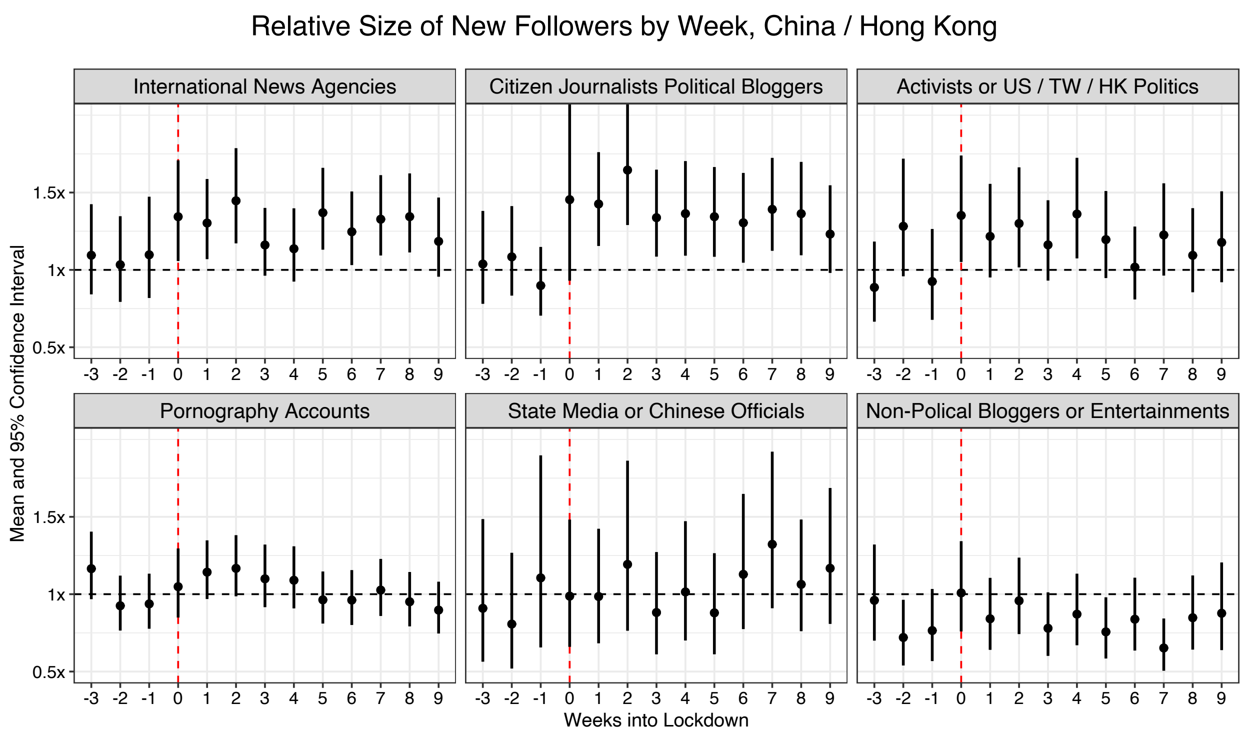
In this section, we assess whether the result is driven by (1) a misspecified treatment period, (2) the choice of comparison group, or (3) an increase of followers due to only a few accounts.

Figure A6 plots the estimates based on regressions for each week before and after the lockdown. We do not see pre-treatment increases in number of followers in China, and the increase starts precisely on the week of lockdown.

Figures A7 and A8 verify that the results in Figure 5 are not due to choosing Hong Kong for the denominator. Figure A7 uses accounts from Taiwan for the denominator, and Figure A8 uses accounts in the United States. These accounts are from any user using Chinese and their self-reported location is in Taiwan or the United States. Figure A9 reports the regression estimate for the relative ratio of number of new followers (akin to a Difference-in-differences design with December 2019 as control period and Hong Kong/Taiwan/China as control group). The result is not driven by Hong Kong-specific trend of news cycles.

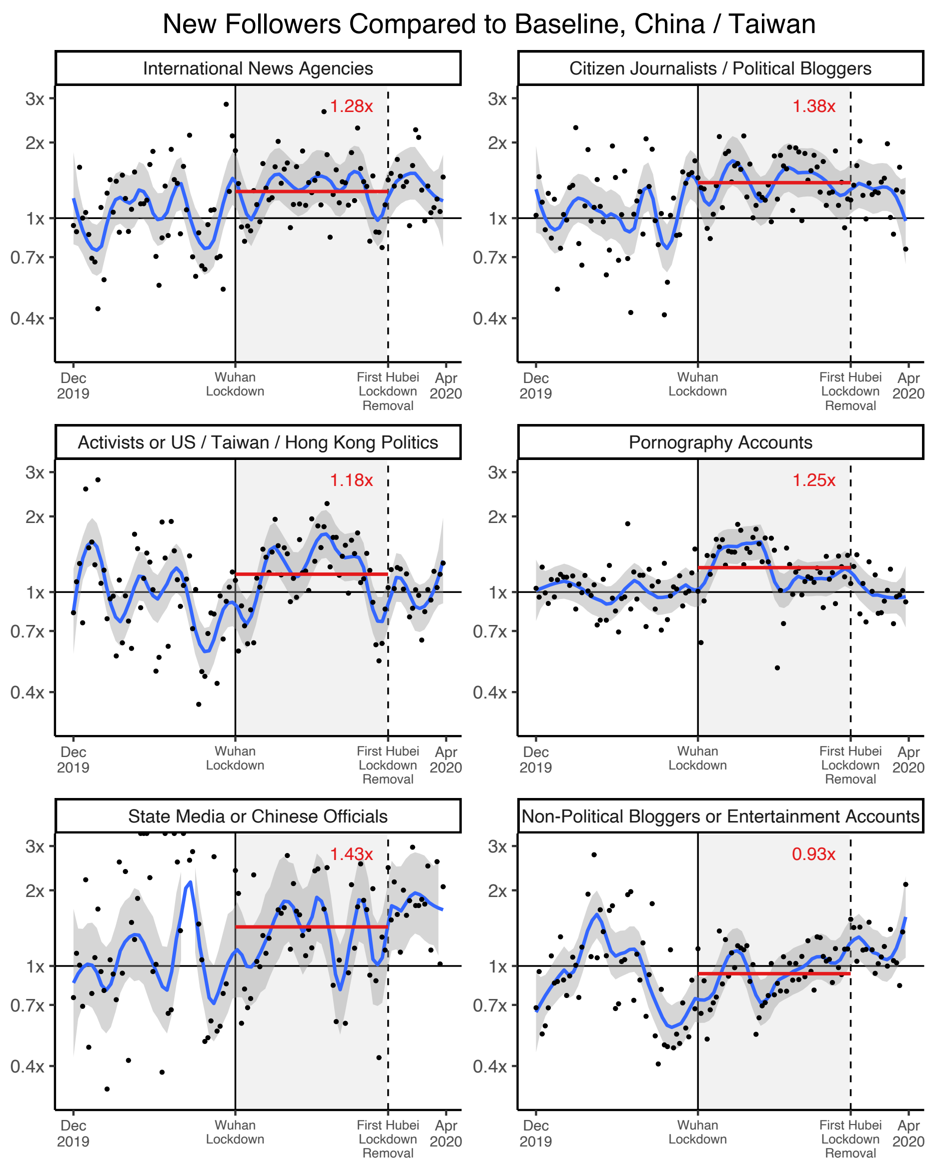
One might also curious about whether new users stayed on Twitter at different rates. Figure A10 plots the daily unique active users since their sign up dates in 2020. We don't find that users from one location stayed on Twitter longer than others.

Figure A6: Increases in Twitter Followers from mainland China versus Hong Kong by Week



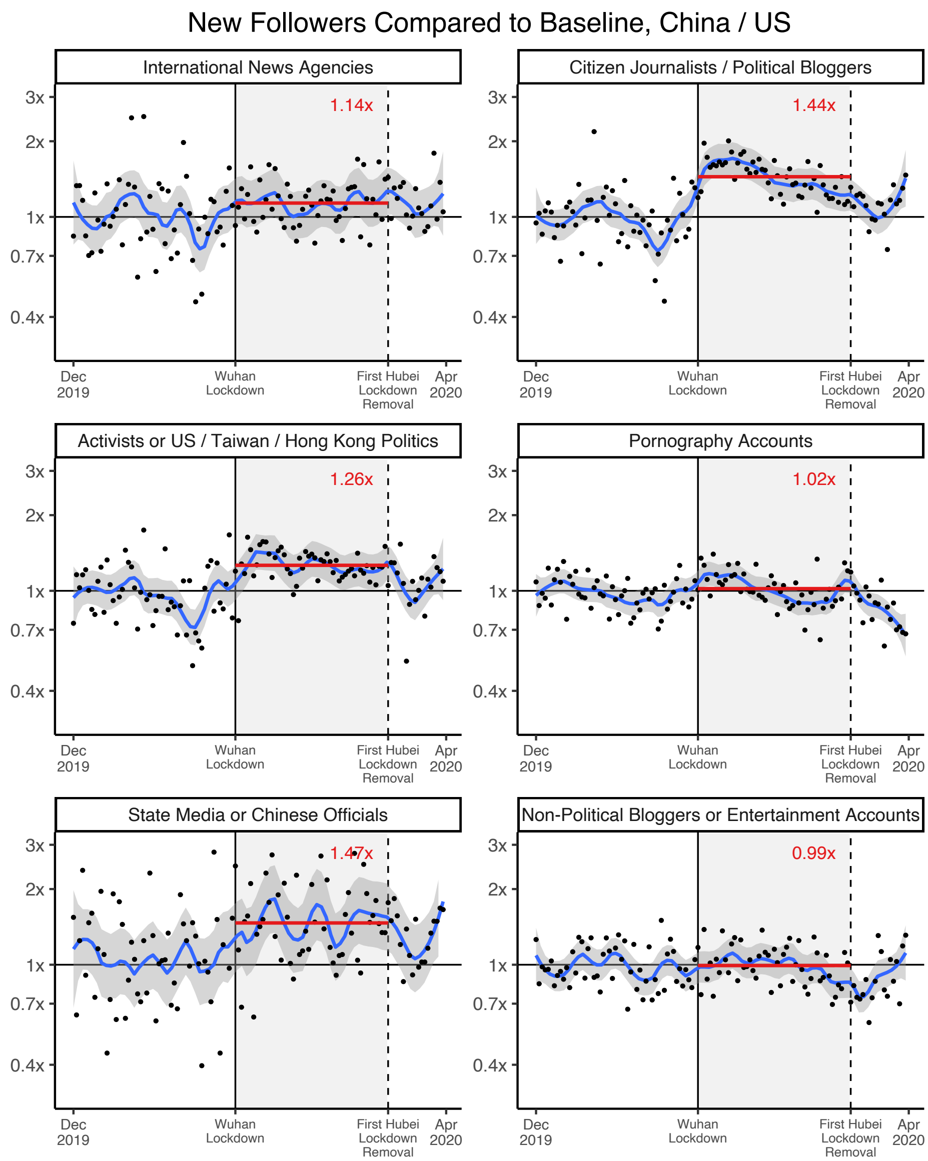
Note: Incidence rate ratios shown above are from Negative Binomial regressions of number of daily new followers on the interaction between dummy for each week and China, with December 2019 as control period and Hong Kong as control group.

Figure A7: Increases in Twitter Followers from China versus Taiwan



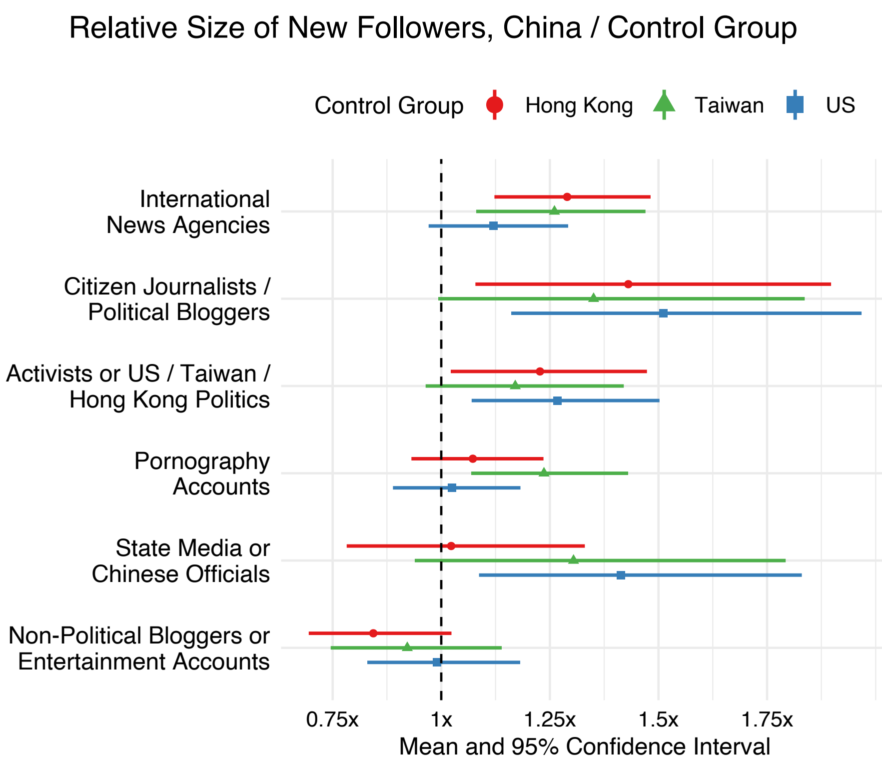
Note: Gain in followers from mainland China compared to Taiwan across six types of popular accounts, relative to December 2019 average. A value greater than 1 means more followers than expected from mainland China than from Taiwan. Accounts creating sensitive, censored information receive more followers than expected once the Wuhan lockdown starts. Fewer Taiwanese users follow Chinese state media or government officials than Hong Kong users do.

Figure A8: Increases in Twitter Followers from China versus US



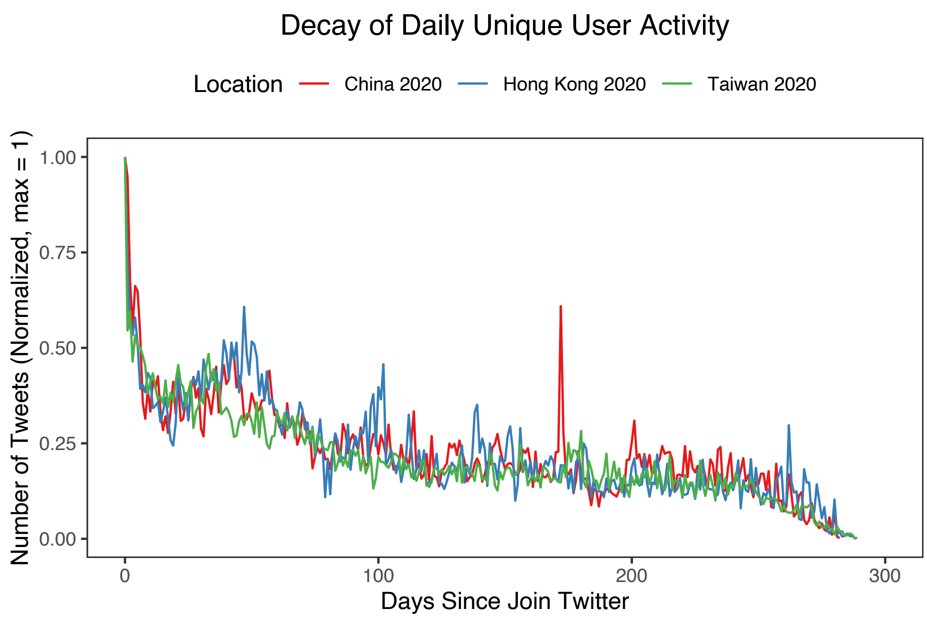
Note: Gain in followers from mainland China compared to US across six types of popular accounts, relative to December 2019 average. A value greater than 1 means more followers than expected from mainland China than from the US. Accounts creating sensitive, censored information receive more followers than expected once the Wuhan lockdown starts. Fewer US users follow Chinese state media or government officials than Hong Kong users do.

Figure A9: Increases in Twitter Followers from China versus Others (Regression Estimate)



Note: Incidence rate ratios shown above are from negative binomial regressions of number of new followers on the interaction between indicator variables for 'in lockdown period' and 'in mainland China', with December 2019 as the control period.

Figure A10: New Users Stay on Twitter at the Same Rates across Locations



Note: This figure plots the daily unique active users since their sign up date using the user panel across locations. A user is considered active between their sign up date and the last day they tweet (before July 2020). We find that users stay on Twitter at the same rate across locations.

6 Wikipedia Country Comparisons

6.1 Page view analysis

Page view data analyzed in this paper is publicly available and hosted here: `https://dumps.wikimedia.org/other/pagecounts-ez/merged/`. In replication materials, we will additionally provide processed and aggregated versions of the page view data so that this paper's findings can be more quickly replicated than would be possible with the above page view files.

Below, we show the top Wikipedia pages by relative and absolute increases in page views within each of the categories we analyzed in the main text, as well as pages about the coronavirus and COVID-19 (pages considered: coronavirus, COVID-19, ventilator, flu, pneumonia, fever). The largest relative increases among these pages and for current leaders were related to coronavirus – the COVID-19 pandemic Wikipedia page and the head of China's National Health Commission. Top increases for pages that were blocked prior to the introduction of https on Wikipedia (after which China blocked all pages) were for an activist who criticized China's pandemic response.

Table A2: Top relative increases for Wikipedia pages January 24 through March 13 compared to December 2019.

Overall	**Blocked**	**Current Leaders**	**Historical Leaders**
马晓伟 _(官员) (36.67) Ma Xiaowei	许志永 (16.78) Xu Zhiyong	马晓伟 _(官员) (36.67) Ma Xiaowei	胡锦涛 (1.81) Hu Jintao
许志永 (16.78) Xu Zhiyong	2 月 17 日 (9.01) February 17	孙春兰 (9.38) Sun Chunlan	邓小平 (1.75) Deng Xiaoping
孙春兰 (9.38) Sun Chunlan	西藏人民起义日 (7.04) Tibetan Uprising Day	李克强 (2.52) Li Keqiang	江泽民 (1.65) Jiang Zemin
2 月 17 日 (9.01) February 17	台湾 (5.21) Taiwan	王岐山 (2.50) Wan Qishan	华国锋 (1.44) Hua Guofeng
西藏人民起义日 (7.04) Tibetan Uprising Day	圆周率日 (4.15) Pi Day	肖捷 (2.45) Xiao Jie	毛泽东 (1.15) Mao Zedong
肺炎 (5.38) Pneumonia	艾未未 (3.93) Ai Weiwei	韩正 (2.14) Han Zheng	
台湾 (5.21) Taiwan	李長春 (3.71) Li Changchun	胡春华 (1.99) Hu Chunhua	
流行性感冒 (5.04) Influenza	新唐人电视台 (3.51) New Tang Dynasty Television	苗圩 (1.88) Miao Wei	
圆周率日 (4.15) Pi Day	唐柏桥 (3.34) Tang Baiqiao	习近平 (1.80) Xi Jinping	
艾未未 (3.93) Ai Weiwei	长春围困战 (3.21) Siege of Changchun	杨晓渡 (1.73) Yang Xiaodu	

Note: This is where authors provide additional information about the data, including whatever notes are needed.

In Figure A11, we show the trajectories for categories matching those analyzed for China –

Table A3: Top absolute *daily* increases for Wikipedia pages January 24 through March 13 compared to December 2019.

Overall	**Blocked**	**Current Leaders**	**Historical Leaders**
Rest of Wikipedia (1095913)	习近平 (4797) Xi Jinping	习近平 (4797) Xi Jinping	江泽民 (1197) Jiang Zemin
2019 冠状病毒病 (new page: 9236) Coronavirus disease 2019	王岐山 (2168) Wang Qishan	王岐山 (2168) Wang Qishan	邓小平 (1102) Deng Xiaoping
习近平 (4797) Xi Jinping	台湾 (2063) Taiwan	李克强 (1584) Li Keqiang	胡锦涛 (1079) Hu Jintao
肺炎 (4603) Pneumonia	六四事件 (1941) June 4 Incident (Tiananmen Square)	孙春兰 (1350) Sun Chunlan	毛泽东 (349) Mao Zedong
流行性感冒 (2463) Influenza	香港电台 (1689) Radio Television Hong Kong	韩正 (579) Han Zheng	华国锋 (255) Hua Guofeng
王岐山 (2168) Wang Qishan	中华人民共和国 (1631) People's Republic of China	胡春华 (541) Hu Chunhua	
台湾 (2063) Taiwan	李克强 (1584) Li Keqiang	马晓伟 _(官员) (244) Ma Xiaowei	
六四事件 (1941) June 4 Incident (Tiananmen Square)	江泽民 (1197) Jiang Zemin	王毅 (119) Wang Yi	
香港电台 (1689) Radio Television Hong Kong	中华民国 (1128) Republic of China	傅政华 (99) Fu Zhenghua	
中华人民共和国 (1631) People's Republic of China	邓小平 (1102) Deng Xiaoping	肖捷 (63) Xiao Jie	

Note: Studying average daily increases standardizes the different lengths of time before versus after the Wuhan lockdown. Labels are limited to: blocked, leader, historical leader, COVID/coronavirus. All other pages are aggregated as "rest of Wikipedia".

current leaders (using offices listed in the CIA World Factbook), historical leaders, and, in Iran, pre-https blocked Wikipedia pages (Nazeri and Anderson 2013).

Russia, Germany, and Italy (none of which block Wikipedia) saw increases in current leader views without accompanying increases in historical leader views. Germany and Italy did see spikes views of in historical leader pages in the weeks leading up to the relaxation of lockdowns in early May, but saw no change during the initial crisis.

German and Russian political pages also saw an increase in political leader page views prior to their own lockdown, and approximately at the same time as the announcement of widespread lockdown in Italy (see Figure A11).

Table A4: Lockdown dates

Country	**Lockdown Start**	**Lockdown End**	**Historical Leaders**
China	January 24, 2020	March 13, 2020	Paramount Leader
	Hubei Lockdown		
Iran	March 20, 2020	April 18, 2020	President, Supreme Leader
	Nowruz - Tehran Easing		
Russia	March 28, 2020	May 12, 2020	President
	Non-Working Period		General Secretary (Soviet Union)
			Chairman, Council of Ministers (1953)
Germany	March 22, 2020	May 6, 2020	Chancellor
	National Social Distancing		
Italy	March 9, 2020	May 18, 2020	Prime Minister
	National Quarantine		

Note: This table lists the time periods we use to estimate the effects of crisis lockdowns on Wikipedia page views, along with the offices considered for the historical leaders analysis. Each country's lockdown involved various levels of lockdown for different parts of the countries, and so there is no single time period for us to analyze. Figure A11 displays Wikipedia page views with solid, vertical gray lines for the periods listed above.

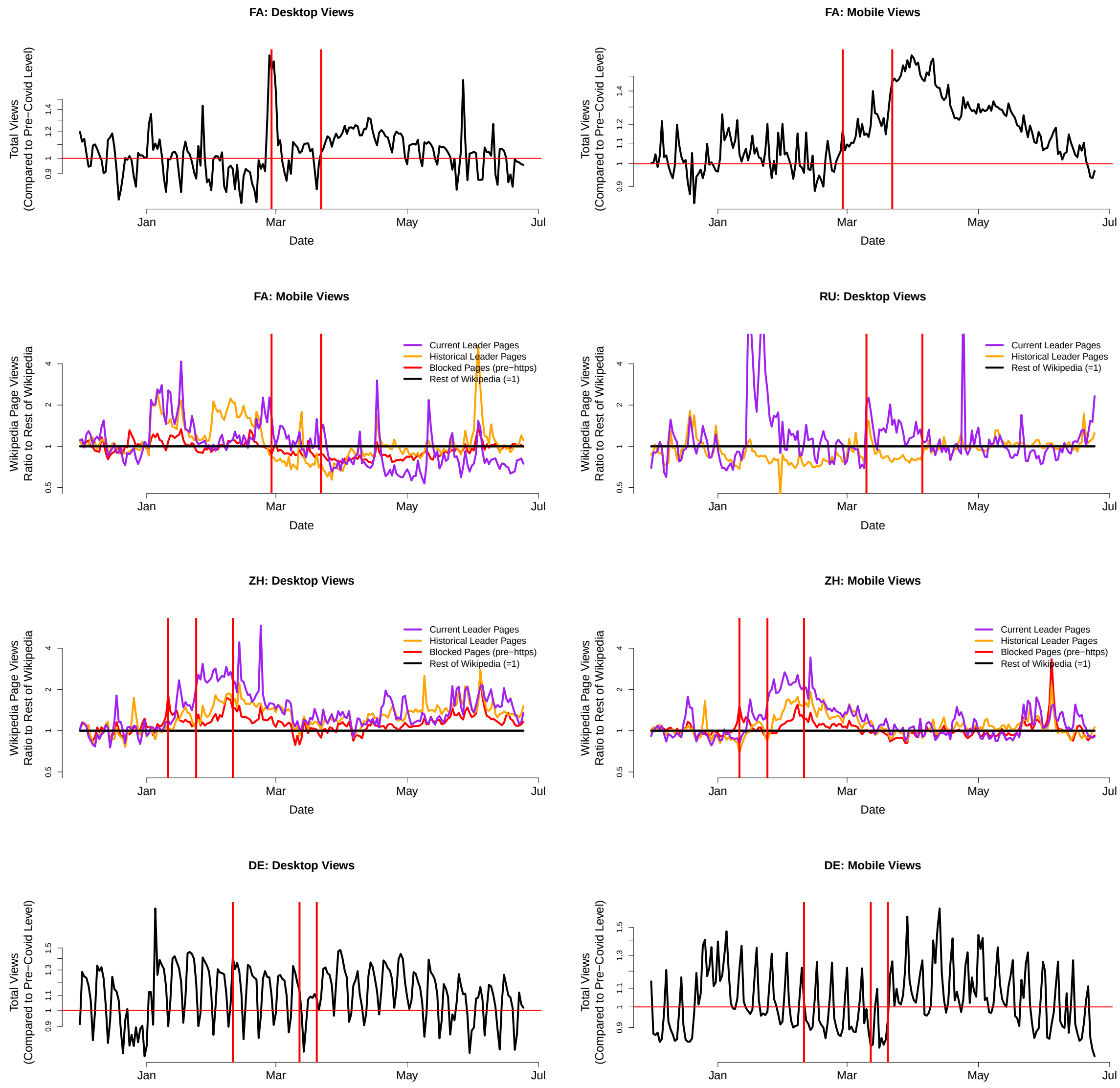


Figure A11: Views of Blocked, Current Leader, and Historical Leader Wikipedia Pages in Other Countries

6.2 Analysis of an expanded set of historical political pages and ‘politically sensitive’ pages using Wikipedia2vec

We replicated our analyses of historical Wikipedia pages and “politically sensitive” (pages specifically blocked in China and Iran prior to the introduction of https) Wikipedia pages by expanding the original set of pages to a much larger set of related pages. We expanded these lists of pages using Wikipedia2vec (Yamada et al. 2020). This analysis assesses 1) whether the increase in views of Chinese historical leaders (and the lack of increase for other languages) was a relatively narrow effect or much broader one than what we see for that small set of pages and 2) whether a broader set of ‘politically sensitive’ pages are able to uncover increases in page views in Iran and Russia. Because, unlike China and Iran, Russia did not provide a list of politically sensitive pages (by blocking specific pages on Wikipedia), we assess Russian views of political opposition pages related to a) Alexei Navalny (arguably the most prominent opposition leader in Russia) and b) a list of opposition-related pages which we mine to discover increases in views – after this, we then looked to closely related pages to assess whether single page increases represented broader trends or were isolated and potentially random occurrences.

Wikipedia2vec finds similar pages (along with other entities and words) on Wikipedia by analyzing the network of page links, the co-occurrence of words, and the occurrences of specific words on pages. This analysis is accomplished using the same approach as in word2vec (Mikolov et al. 2013). At a high level, this approach involves placing words and entities into a shared n-dimensional space such that words and entities are placed closed together if they frequently share contexts (e.g. page links or co-occurring words). Shared contexts must occur beyond what would be expected from the frequency of a word or page, which is accomplished through ‘negative sampling’ – predicting the co-occurrence of words and entities against frequency weighted sampling of negative cases. Once in the n-dimensional space, we can find the most similar entities (pages) for any given entity (or the mean of a set of entities’ projections) using cosine similarity – and we can incorporate dissimilar entities in this calculation by flipping the sign those entities’ locations when calculation the mean of a set of entities. Wikipedia2vec can be run with hyperparameters that affect the size of the n-dimensional space and the exact weighting scheme used in negative sampling. Our estimations for each language used the same default settings as the wikipedia2vec pre-trained embeddings provided at `https://wikipedia2vec.github.io/wikipedia2vec/pretrained/` with the number of dimensions set to 100.

For each set of pages (historical leaders, blocked pages, current leaders, Russian opposition pages), we found the top 100, 250, 500, and 1,000 pages that were most similar according to Wikipedia2vec. For historical leaders, we expanded to historical leader related pages not related to the current leader – using the current leader as a dissimilar case – and we expanded to current leader related pages not related historical leaders. With these sets, we re-estimated the changes in views during the first 30 days of lockdown. This excludes the late lockdown spikes in historical leader views visible for German and Italian (visible in Figure 7 in the main text and in Figure A11 above). Note that the German increases in views of historical leaders (in those figures and in the results below) began well prior to the German lockdown (in February).

For Alexei Navalny specifically, we also manually collected a list of Wikipedia pages closely related to his opposition activities, and re-estimated changes in lockdown for each of these pages. The list of Russian opposition-related pages checked for increases is shown in Table A5.

For the pages previously blocked in Iran, Nazeri and Anderson (2013) provided labels for the category of each blocked page: academic, artistic/cultural, drugs or alcohol, human rights, media and journalism, other, political, profane non-sexual, religious, and sex and sexuality. We also replicated our analyses for the Persian language set subset to page categories human rights, media and journalism, and political.

The findings from these analyses are displayed in Figures A12, A13, and A14, and we also show findings for current leaders in Figure A15. In each cluster of estimates in the top panels, the first is the estimate for the seed pages (and is colored yellow for historical pages, red for blocked/'politically sensitive' pages, purple for current leaders). These exclude estimates for seeds which we *mined* for increases (i.e. selected them *only* because we saw increases during lockdowns – after a Bonferroni multiple testing correction). Given many tests when looking for increases, these pages have estimates that could very likely reflect random variation in page views, even though we are relatively certain that the increases were not *zero*, given the multiple testing correction.

Across the results, we see 1) that the increase in historical leader page views in Chinese also applies to a much larger set of pages and page views (bottom panel) and 2) we do not see comparable increases in historical leader pages or politically sensitive page views in other languages, despite increased interest in *current* leaders across almost all languages analyzed.

In the manual Alexei Navalny analysis, we see that views for his page specifically did rise and that this rise was comparable to what we see for historical leaders in Chinese. However, unlike the broad increase in views in China, we did not see similar increases for any other Navalny-related pages – and only one of the 9 considered showed a statistically significant increase without a multiple-testing correction (falling just short of significance at a 0.05 level after a Bonferroni correction for 9 tests).

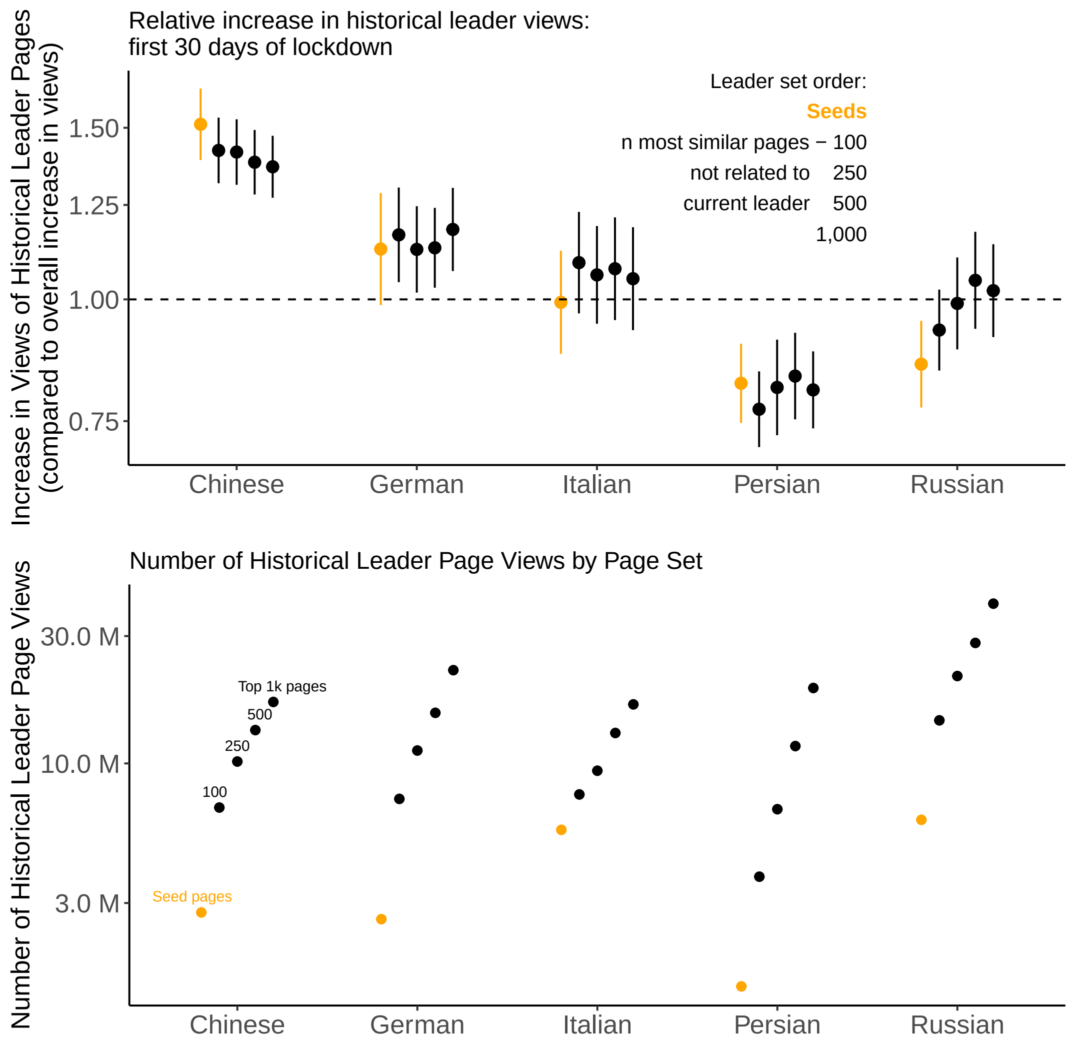


Figure A12: *Changes in views of historical leader Wikipedia pages (expanded set of pages).* German increases in views of historical leaders began in February (see Figure A11 above)

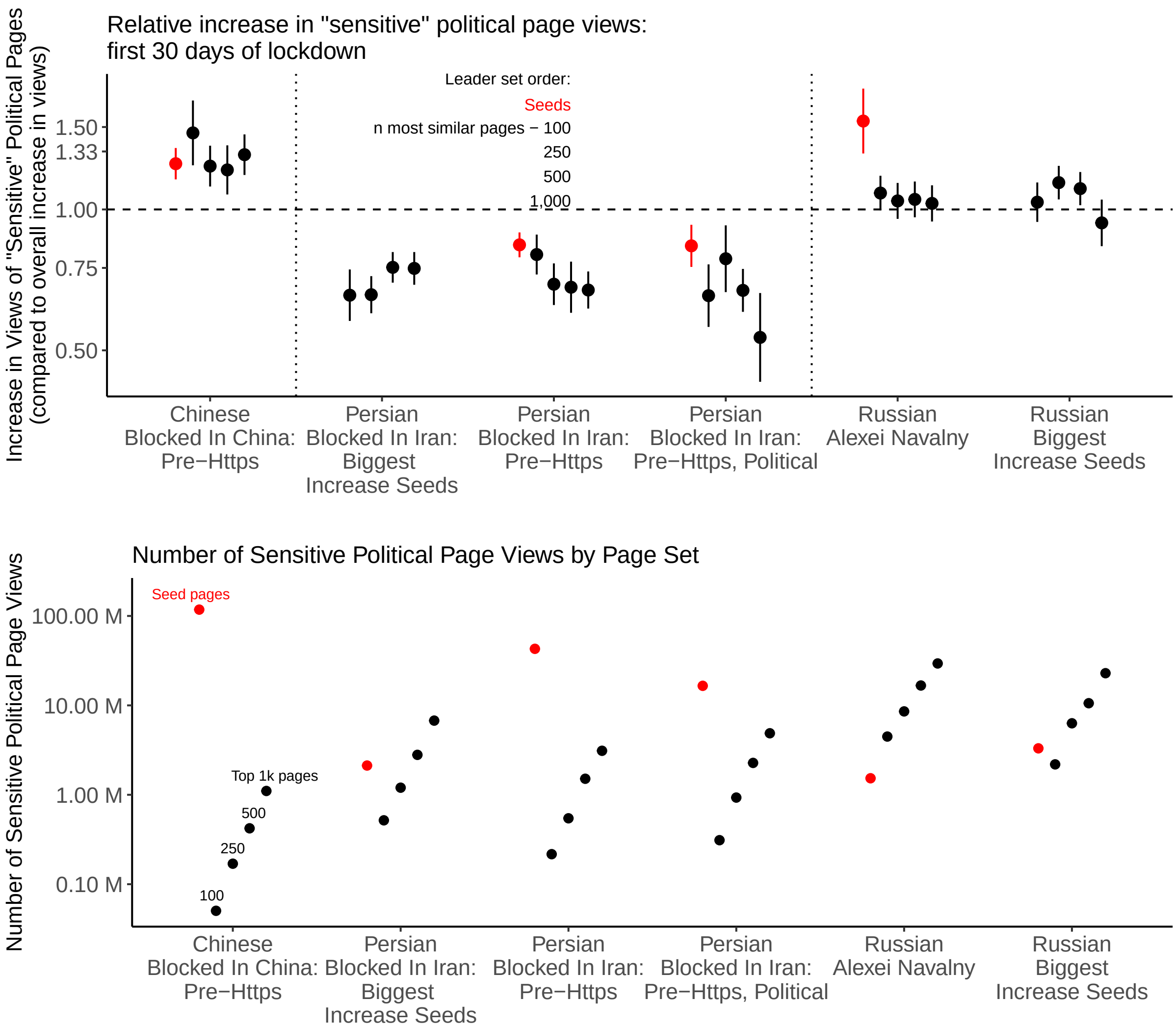


Figure A13: *Changes in views of 'politically sensitive' Wikipedia pages (expanded set of pages).*

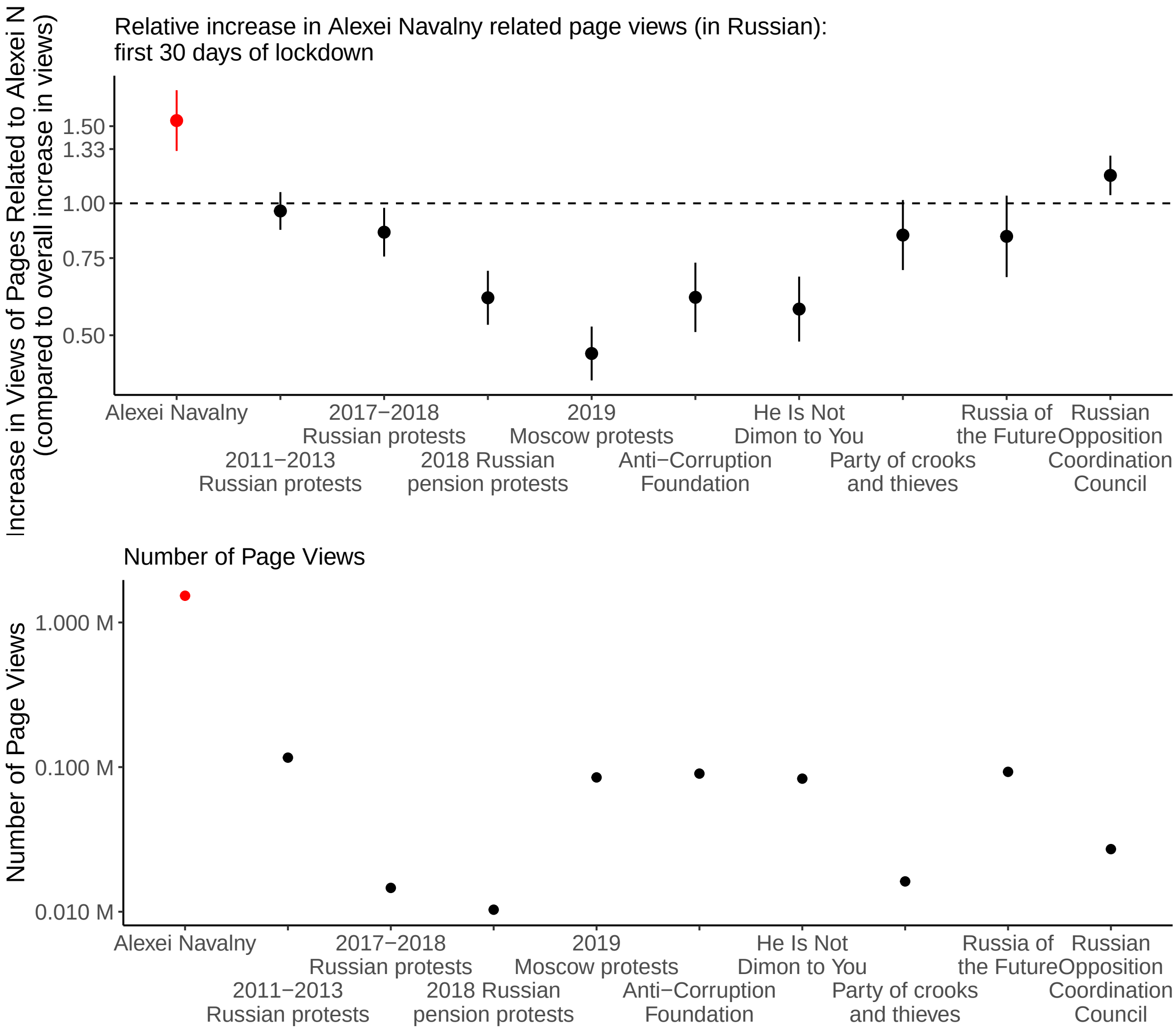


Figure A14: *Changes in views of Alexei Navalny related Wikipedia pages.* The Alexei Navalny-related pages in this figure are listed in alphabetical order.

Table A5: List of opposition-related pages in Russian that were checked for significant increases during lockdown.

2011-2013 Russian protests	He Is Not Dimon to You
2014 anti-war protests in Russia	Human rights in Russia
2017-2018 Russian protests	List of journalists killed in Russia
2018 Russian pension protests	Media freedom in Russia
2019 Moscow protests	**Mikhail Khodorkovsky**
Alexander Litvinenko	Novaya Gazeta
Alexei Navalny	**Open Russia**
Anna Politkovskaya	**Opposition to Vladimir Putin in Russia**
Anti-Corruption Foundation	Party of crooks and thieves
Assassination of Anna Politkovskaya	Pussy Riot
Assassination of Boris Nemtsov	Russia of the Future
Boris Berezovsky (businessman)	Russian Opposition Coordination Council
Boris Nemtsov	Sergei Magnitsky
Corruption in Russia	Sergei Yushenkov

Note: Russia did not block specific Wikipedia pages prior to Wikipedia's introduction of https. Because of this, we do not have a government-provided list of politically sensitive or objectionable content. As an alternative, we mine a manual list of government opposition-related pages, and then check whether for those increases were narrow and perhaps random (i.e. only occurred for those specific pages) *or* represented broad increases similar to those seen for historical and previously blocked pages in China. This table lists those Wikipedia pages (translated) that were checked for significant associations during the Russian lockdown period when compared to December 2019. Pages with statistically significant increases ($p < 0.05$) after a Bonferroni multiple testing were used as seeds when expanding with Wikipedia2vec. These "biggest increase seeds" are in bold above.

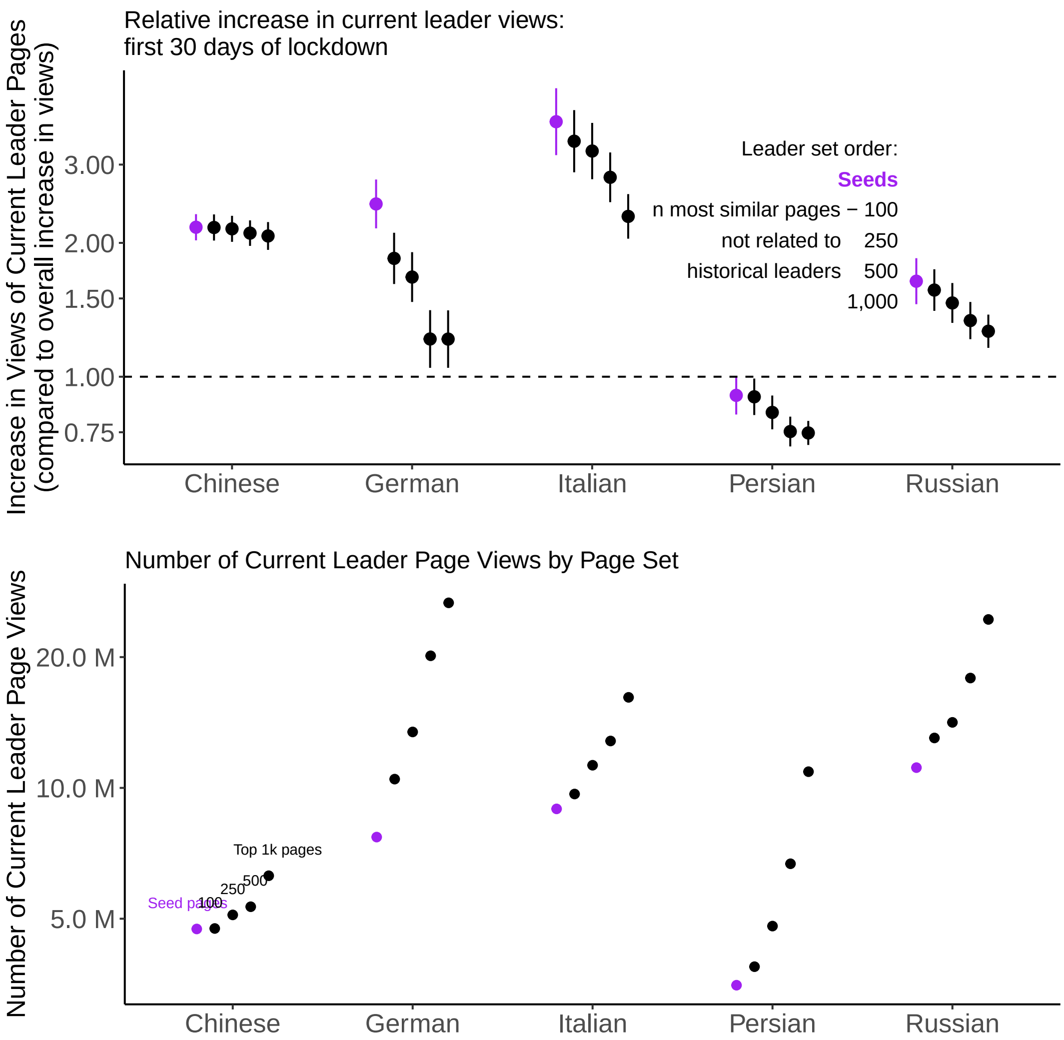


Figure A15: *Changes in views of current leader Wikipedia pages (expanded set of pages).*

7 Text Analysis of Tweets

7.1 Hand labels

To better understand the range of political content that Chinese social media users may have encountered on Twitter, we hand labeled a random sample ($n = 500$) of tweets that 1) mentioned China or Chinese provinces and 2) mentioned any country [1] between December 1st, 2019 and December 1st, 2020.

We evaluate the tone of each tweet (positive, negative, or neutral) and the main country being referred to in the tweet. These labels primarily evaluate whether to what extent Chinese social media users may have been exposed to negative content about China *and* negative content about other countries, including other countries' pandemic responses. However, we also label the main topic of the tweet (reading the text in Chinese to identify what appear to be the most frequent topics). They include the handling of COVID in various countries, Hong Kong protests, US elections, and other recent developments between US-China relations. We provide a comparison to these topics using the output of an automated topic model below.

Overall, we find that the account types that saw a disproportionate increase in followers at the start of the pandemic tended to cover *both* China and other countries in a negative way (Figure A16). Coverage of both US and China are mostly negative in international news agencies, while in Chinese state media or Chinese officials (which did not see the same increase in followers) the coverage of China is strictly positive and the coverage of US is only negative. Furthermore, we also find that the tone on different topics was dissimilar between different types of popular accounts (Figure A17). For international news agencies and activists/citizen journalists, tweets about the U.S. election, COVID in the U.S., Hong Kong protests, and U.S.-China disputes are all relatively negative. In tweets by Chinese state media or officials, the coverage about COVID in China and Chinese economic development are positive.

Notably, the United States was mentioned much more than other countries, suggesting that *general* international comparisons on COVID-19 pandemic responses might have been relatively rare compared to content about disputes with the United States and, perhaps, with the U.S. presidential administration.

[1] We identified country mentions using the Unicode Common Locale Data Repository: `https://www.unicode.org/Public/cldr/39/` (Chinese language country names in core/common/main)

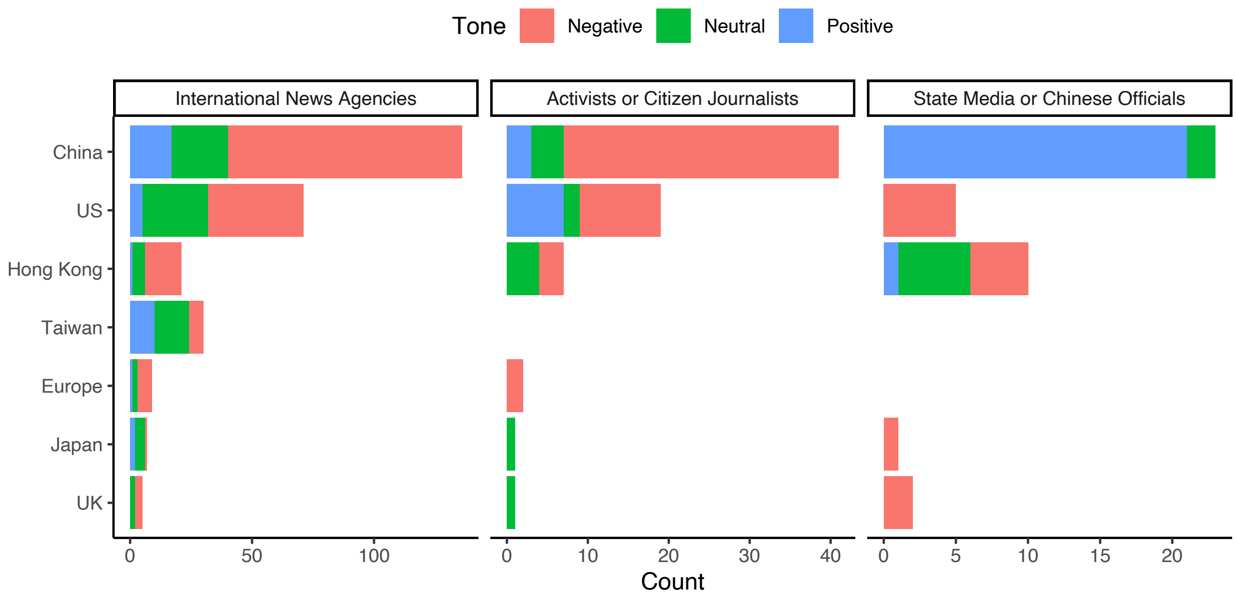


Figure A16: Tone of tweets by popular Twitter accounts across main countries mentioned

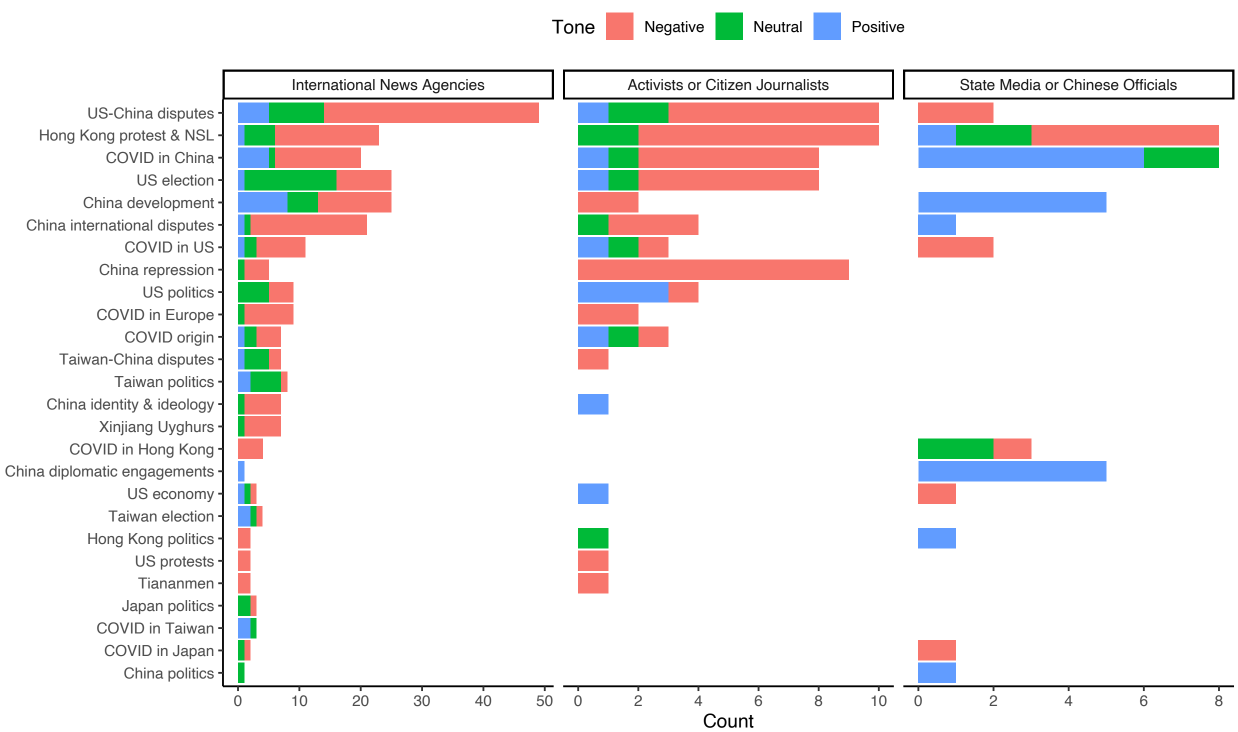


Figure A17: Tone of tweets by popular Twitter accounts across topics of tweets

7.2 Topic models

To supplement the hand labels, we ran a topic model on tweets that mentioned China (Table A6) or any country (Table A7). In combination with the hand labels, this evaluates whether coverage of China was uniformly negative or more mixed, and also whether other countries and/or their disputes with China may have been covered in ways potentially favorable to the Chinese government and Communist Party. In international comparisons, for example, we might find favorable comparisons in other countries' handling of the COVID pandemic, coverage of international disputes thought to drive nationalist sentiment in China, or news on anti-Asian racism in the United States and elsewhere.

These topic models were run on 10 thousand tweet samples. Like the hand labels, these simple random samples were drawn from tweets that mentioned China or Chinese provinces (or any country for the second model) in Chinese from December 2019 through December 6, 2020[2]. We also restricted the samples to content from the account categories "International News Agencies", "Citizen Journalists / Political Bloggers", and "Activists or US / Taiwan / Hong Kong Politics" – the categories which saw increases in views during the pandemic, and there were both political and not associated with Chinese state media or officials.

Each of topic models was estimated using the structural topic model R package (Roberts et al. 2016) [3] with the number of topics set to 50, and the structural topic models were estimated without including covariates.

Several topics are related to politically sensitive issues in China (e.g. Tiananmen Square, the Hong Kong national security law, Xinjiang and human rights), but also many topics are related to international disputes with China (especially between China and the United States/President Trump) and the COVID-19 pandemic around the world. We hope this exploratory analysis will help guide future work.

[2]This slightly differs from the December 1 period above because the hand label and topic model analyses were conducted separately by different members of the research team. We decided that a closer alignment in the time frames would have no meaningful influence on these exploratory analyses.

[3]https://www.structuraltopicmodel.com/

Table A6: Topic model on tweets that mention China.

Topic label	Highest probability keywords	Proportion of corpus
US-China trade	美国, 中国, 美, 普, 朗, 奥, 中, 佩, 贸易, 总统 America, China, Trump, Austria, China, trade, President	0.07
Chinese economy	中国, 经济, 公司, 亿, 企业, 美元, 全球, 市场, 投资, 商 China, economy, company, billion, enterprise, USD, global, market, investment, business	0.05
COVID in China (1)	疫, 情, 病毒, 中国, 新冠, 武汉, 卫, 肺炎, 冠状, 爆发 epidemic, emotion, virus, China, COVID-19, Wuhan, health, pneumonia, coronavirus, outbreak	0.05
Chinese media	中国, 媒体, 记者, 报道, 推, 信, 新闻, 网络, 微, 媒 China, media, reporters, reports, tweets, letters, news, internet, micro, media	0.04
HK National Security Law	香港, 法, 时间, 国安, 港, 英国, 日, 版, 会议, 北京 Hong Kong, France, time, national security, Hong Kong, UK, Japan, edition, conference, Beijing	0.03
COVID in China (2)	病例, 诊, 确, 日, 肺炎, 感染, 死亡, 例, 新增, 湖北 case, diagnosis, confirmation, day, pneumonia, infection, death, case, new, Hubei	0.03
Tiananmen Square protests	自由, 政府, 中国, 政治, 批评, 四, 言论, 人士, 行动, 六 freedom, government, China, politics, criticism, four (June 4 incident), speech, people, action, six	0.03
Xi Jinping & Donald Trump	习, 近平, 川, 普, 新闻, 马, 陈, 今天, 李, 中国 Xi Jinping, Trump, news, Ma, Chen, today, Li, China	0.03
China's international disputes	中国, 印度, 军, 海, 日本, 南, 军事, 机, 中, 冲突 China, India, army, sea, Japan, South, military, aircraft, China, conflict	0.03
Human rights in China (1)	维, 权, 律师, 人士, 遭, 罪, 警方, 案, 公民, 当局 protection, rights, lawyers, persons, victims, crimes, police, cases, citizens, authorities	0.03
Chinese Communist Party	中共, 人民, 政权, 世界, 中国, 反, 民众, 统治, 共, 内部 CCP, people, regime, world, China, anti, people, domination, communist, internal	0.03
COVID in China (3)	中國, 武漢, 中共, 美國, 習, 新聞, 與, 對, 診, 確 China, Wuhan, CPC, United States, Xi, news, and, right, medical, confirmation	0.03
COVID in China (4)	武汉, 封, 肺炎, 城, 医院, 隔离, 湖北, 人员, 疫, 感染 Wuhan, closure, pneumonia, city, hospital, quarantine, Hubei, personnel, epidemic, infection	0.03
Taiwan-China relationship (1)	党, 台湾, 共产, 民主, 中国, 台, 人民, 蔡, 英文, 代表 Party, Taiwan, Communist, democracy, China, Taiwan, people, Tsai Ing-wen, English, representative	0.03
Chinese people	中国人, 现在, 应该, 没有, 很多, 这种, 都是, 知道, 没, 世界 Chinese, now, should, no, a lot, this kind, all, know, no, world	0.03
Human rights in China (2)	国, 中国, 人权, 国际, 联合, 组织, 声明, 国家, 中华, 世界 country, China, human rights, international, union, organization, statement, country, China, world	0.02
Chinese diplomacy	中国, 希望, 德国, 欧洲, 欧盟, 澳洲, 毅, 王, 国家, 一个 China, hope, Germany, Europe, EU, Australia, Wang Yi, country, one	0.02
China's medical aid	中国, 口罩, 澳大利亚, 文化, 国家, 一周, 日本, 援助, 医疗, 法国 China, mask, Australia, culture, country, one week, Japan, aid, medical, France	0.02
US-China disputes (1)	中国, 美国, 局, 签证, 调查, 州, 政府, 间谍, 情报, 名 China, United States, bureau, visa, investigation, state, government, spy, intelligence, name	0.02
	中国人, 一个, 吃, 思想, 没有, 里, 国家, 中国, 社会, 你的	0.02

Chinese society	Chinese, one, eat, thought, no, inside, country, China, society, yours	
Xinjiang internment/’reeducation’ camps	新疆, 维吾尔, 人权, 营, 民族, 族, 教育, 穆斯林, 集中, 宗教 Xinjiang, Uyghur, human rights, camp, ethnic, ethnic, education, Muslim, concentration, religion	0.02
Chinese education	大学, 学生, 教授, 学院, 中国, 留学生, 蒙古, 教育, 学, 万 university, student, professor, college, China, international student, Mongolia, education, learning, million	0.02
China’s economic development	中国, 问题, 视, 国家, 认为, 更, 发展, 春, 研究, 库 China, issues, view, country, think, more, development, spring, research, library	0.02
Shanghai	上海, 张, 店, 一个, 冯, 家, 朋友, 虎, 买, 民 Shanghai, Zhang, shop, one, Feng, home, friends, tiger, buy, people	0.02
Chinese politics	长, 副, 委, 强, 中央, 书记, 马, 组, 王, 志 Chief, Deputy, committee, strong, Central, Secretary, Ma, group, Wang, Zhi	0.02
China’s online patriotism	网, 民, 主义, 联, 中国, 互, 爱国, 爱, 中, 一个 net, people, doctrine, union, China, mutual, patriotic, love, China, one	0.02
Wang Quanzhang	王, 警察, 全, 北京, 璋, 车, 派出所, 文, 黎, 李 Wang Quanzhang, police, full, Beijing, car, police station, Li, Li	0.01
Taiwan-China relationship (2)	北京, 离开, 台北, 台灣, 中共, 五, 破, 高层, 日, 空 Beijing, departure, Taipei, Taiwan, CCP, broken, high-rise, day, empty	0.01
Li Wenliang	文, 李, 医生, 亮, 发, 中国, 福建, 吹哨, 医, 找 Li Wenliang, doctor, bright, hair, China, Fujian, whistling, medical, find	0.01
Shandong	山东, 杨, 家, 喜, 丁, 母, 监狱, 号, 戴, 省 Shandong, Yang, home, Ding, mother, prison, number, Dai, province	0.01
China stability	中国, 进入, 不断, 稳, 更加, 维, 度, 期, 月, 更 China, enter, continue, stabilize, more, dimension, degree, period, month, more	0.01
Hubei	黄, 一个, 湖北, 中, 琦, 锋, 写, 先生, 儿子, 两 Huang, one, Hubei, middle, Qi, Feng, write, Mr, son, two	0.01
Africa-China relationship	中国, 非洲, 猪, 粮食, 危机, 到底, 国家, 猪肉, 大豆, 粮 China, Africa, pig, food, crisis, end, country, pork, soy, grain	0.01
Vaccines and China	疫苗, 中国, 实验, 试验, 巴西, 新冠, 测试, 使用, 室, 今日 vaccine, China, experiment, trial, Brazil, COVID-19, test, use, laboratory, today	0.01
Guangdong	广东, 太, 带, 场, 云, 私人, 中, 钱, 市, 镇 Guangdong, Tai, with, field, cloud, private, middle, money, city, town	0.01
China flood	洪水, 洪, 湖北, 江西, 灾, 村, 暴雨, 三峡, 安徽, 南方 flood, flood, Hubei, Jiangxi, disaster, village, heavy rain, Three Gorges, Anhui, south	0.01
Chongqing	重庆, 安, 市, 庆, 工程, 街头, 一个, 两, 桥, 万 Chongqing, An, city, Qing, project, street, one, two, bridge, million	0.01
Xu Zhangrun & Chen Qiushi	许, 志, 实, 章, 失, 润, 强, 秋, 陈, 男 Xu Zhangrun, Zhi, actual, chapter, lost, run, strong, autumn, Chen Qiushi, male	0.01
North Korea	出, 朝鲜, 恩, 传, 中国, 中, 金正, 瑞, 师, 金 out, North Korea, En, biography, China, China, Kim Jong-un, Rui, Division, Kim	0.01
Sichuan earthquake	四川, 成都, 地震, 教会, 家庭, 怡, 牧师, 网, 线, 日 Sichuan, Chengdu, earthquake, church, family, pastor, net, line, day	0.01

China legal system	获, 高, 法官, 律师, 法院, 审, 权, 智, 企业, 晟 gain, high, judge, lawyer, court, trial, right, wisdom, enterprise, Sheng	0.01
Tibet & the Dalai Lama	西藏, 宗教, 奖, 喇嘛, 藏人, 藏, 达赖, 波, 拉, 流亡 Tibet, religion, award, Lama, Tibetan, Tibet, Dalai Lama, Poland, exile	0.01
Henan	周, 郭, 江, 河南, 脸, 宝, 勇, 洋, 上街, 克 Zhou, Guo, Jiang, Henan, face, Bao, Yong, Yang, Going to the steet	0.01
US-China disputes (2)	馆, 领, 关闭, 领事, 驻, 中的, 休, 中, 图书, 洛杉矶 library, collar, closing, consul, consulate, Chinese, close, Chinese, books, Los Angeles	0.01
Falun Gong	复, 工, 非常, 功, 法轮, 迫害, 学员, 分, 吴, 中共 recovery, Falun Gong, extraordinary, persecution, practitioners, points, Wu, CCP	0.01
Overseas Chinese	中国, 没有, 亚太, 已经, 视频, 政府, 接受, 大陆, 海外, 一个 China, no, Asia Pacific, already, video, government, accepted, Mainland, Overseas, one	0.01
Tianjin	成, 天津, 种, 盼, 历史, 肉, 两, 权, 元, 红 Cheng, Tianjin, kind, hope, history, meat, two, Quan, Yuan, red	0.01
the Vatican	引起, 中国, 患者, 续, 关注, 中, 梵蒂冈, 隔离, 协议, 治疗 arouse, China, patient, continued, attention, China, Vatican, quarantine, agreement, treatment	0.01
China's railroads	路, 广州, 捷克, 罕见, 深圳, 返回, 移动, 大会, 日, 铁 road, Guangzhou, Czech Republic, rare, Shenzhen, return, mobile, convention, Japan, iron	0.01
US-China disputes (3)	抵制, 北京, 中共, 中, 人權, 美, 日, 奥, 洁, 佩 boycott, Beijing, CCP, China, human rights, U.S., Japan, Austria, Yang Jiechi, Mike Pompeo	0.01

Table A7: Topic model on tweets that mention any country.

Topic label	Highest probability keywords	Proportion of corpus
China's diplomacy	中国, 政府, 问题, 北京, 政策, 认为, 关系, 官员, 称, 全球 China, government, issues, Beijing, policy, think, relationship, officials, scale, global	0.05
Hong Kong National Security Law	香港, 法, 港, 国安, 中, 版, 民主, 送, 派, 港人 Hong Kong, law, Hong Kong, national security, China, version, democracy, send, send, Hong Kong people	
COVID	疫, 病毒, 情, 武汉, 新冠, 肺炎, 冠状, 卫生, 卫, 世界 epidemic, virus, emotion, Wuhan, novel coronavirus, pneumonia, coronavirus, health, health, world	0.04
China security	公司, 华, 协议, 机构, 安全, 报告, 政府, 中, 提供, 部 company, China, agreement, agency, security, report, government, China, provision, ministry	0.04
US & Taiwan & China	中國, 中共, 美國, 世界, 對, 武漢, 台灣, 與, 將, 國家 China, Chinese Communist Party, the United States, the world, right, Wuhan, Taiwan, and, will, countries	0.03
Chinese economy	经济, 亿, 美元, 万, 企业, 市场, 投资, 银行, 金融, 公司 economy, billion, USD, million, enterprise, market, investment, bank, finance, company	0.03
China human rights (1)	国际, 国, 人权, 联合, 组织, 声明, 欧盟, 日, 调查, 呼吁 international, national, human rights, union, organization, statement, EU, Japan, investigation, appeal	0.03
Mike Pompeo	中共, 美, 奥, 佩, 蓬, 馆, 制裁, 中, 国务卿, 中美 CPC, United States, Mike Pompeo, Pei, Peng, pavilion, sanctions, China, Secretary of State, Sino-US	0.03
Communism	社会, 主义, 世界, 反, 西方, 化, 种族, 历史, 共产, 一个 society, doctrine, world, anti, Western, transformation, race, history, communism, one	0.03
US election	党, 民主, 选举, 大选, 议员, 共产, 州, 投票, 共和, 两 Party, democracy, election, general election, congressman, communist, state, vote, republic, two	0.03
Taiwan election	台湾, 台, 英文, 蔡, 军, 美, 大陆, 总统, 日, 中 Taiwan, Taiwan, Tsai Ing-wen, military, United States, Mainland China, president, Japanese, Chinese	0.03
COVID	日, 诊, 确, 病例, 死亡, 感染, 例, 人数, 新增, 新冠 day, diagnosis, confirmed, case, death, infection, case, number, new, novel coronavirus	0.03
Weibo	媒体, 推, 视频, 文, 发, 信, 微, 信息, 纽约, 博 media, Twitter, video, text, post, letter, Weibo, information, New York, blog	0.02
Chinese immigrants	现在, 一个, 没有, 都是, 很多, 生活, 里, 华人, 移民, 想 now, one, none, all, many, life, Chinese, immigrants, thinking	0.02
Trump	普, 川, 总统, 朗, 白宫, 支持, 表示, 顾问, 日, 竞选 Trump, president, Lang, White House, support, show, advisor, Japan, election	0.02
China human rights (2)	律师, 权, 维, 中国, 当局, 公民, 刘, 许, 人士, 王 lawyers, rights, maintenance, China, authorities, citizens, Liu, Xu, personalities, Wang	0.02
Canada-China trade	宣布, 限制, 加拿大, 贸易, 产品, 制造, 出口, 禁止, 关税, 晚 announce, restrict, Canada, trade, products, manufacturing, export, prohibition, tariff, late	0.02
Xi Jinping	习, 近平, 陈, 时间, 黄, 新闻, 北京, 今天, 中国, 节目 Xi Jinping, Chen, time, Huang, news, Beijing, today, China, programs	0.02
Russia	斯, 俄罗斯, 德, 马, 尼, 利, 罗, 纳, 克, 部长 Slovakia, Russia, Germany, Malaysia, Nigeria, Li, Romania, Croatia, Minister	0.02

US human rights	美国, 已经, 点, 线, 接受, 公民, 人的, 清楚, 不能, 迫使 America, already, point, line, accept, citizen, human, clear, cannot, forced	0.02
COVID in France	法国, 封, 隔离, 航, 医院, 班, 人员, 城, 名, 两 France, seal, quarantine, aviation, hospital, class, personnel, city, name, two	0.02
COVID in UK	英国, 疫苗, 英, 新冠, 德, 约翰, 逊, 政府, 瑞, 药物 UK, vaccine, UK, COVID-19, Germany, Boris Johnson, son, government, Switzerland, drugs	0.02
Chinese news reports	记者, 新闻, 月, 媒, 报道, 传, 林, 名, 报, 官 reporter, news, month, media, report, biography, Lin, name, newspaper, official	0.02
Tiananmen Square protests	自由, 中国人, 四, 不会, 六, 大学, 纪念, 学生, 言论, 周年 freedom, Chinese, four (June 4 incident), no, six, university, commemoration, student, speech, anniversary	0.02
Chinese factories, return to work	出现, 复, 工, 州, 中国, 状态, 症状, 城, 危机, 区 appearance, recovery, work, state, China, state, symptom, city, crisis, district	0.02
China's diplomacy	国家, 一个, 非洲, 利益, 亚, 中, 澳洲, 女性, 发展, 已经 country, one, Africa, interest, Asia, China, Australia, female, development, already	0.02
The Epoch Times	人民, 全世界, 看到, 爆, 支持, 紀元, 平台, 完整, 訂閱, 正義 people, worldwide, see, burst, support, The Epoch Times, platform, complete, subscription, justice	0.02
Japan & China	研究, 日本, 事, 科学, 日, 中, 大学, 恩, 发生, 蒙古族 research, Japan, events, science, Japan, China, university, En, occurrence, Mongolian	0.02
Black Lives Matter	没, 做, 全, 老, 黑, 命, 中国, 真, 文革, 骂 no, do, full, old, black, life, China, true, Cultural Revolution, curse	0.02
US election	拜, 总统, 大选, 候选, 福, 辩论, 副总统, 当选, 团队, 提名 Biden, president, general election, candidate, fortune, debate, Vice President, elected, team, nomination	0.02
Iran & Soleimani	伊朗, 莱, 核, 导弹, 至少, 美军, 袭击, 苏, 曼, 尼 Iran, Lebanon, nuclear, missile, at least, U.S. military, attack, Soviet, Soleimani, Nepal	0.02
China academia	谈, 中国, 实, 大学, 教授, 学者, 文件, 时事, 出版, 之音 talk, China, reality, university, professor, scholar, document, current affairs, publication, voice	0.01
Chinese history	世界上, 你的, 鸡, 一个, 没有, 两, 历史, 中, 清, 地方 in the world, your, chicken, one, none, two, history, middle, Qing, place	0.01
China's internet	网, 联, 民, 中国, 视, 功, 红, 失, 一定, 互 net, united, people, China, vision, power, red, loss, certain, mutual	0.01
Abe & Japan	日本, 一周, 热门, 东京, 安倍, 回顾, 旅游, 抵制, 亚洲, 热点 Japan, week, popular, Tokyo, Abe, review, travel, boycott, Asia, hot	0.01
Wolf warrior diplomacy	欧洲, 留学生, 战, 女, 狼, 新西兰, 国, 总理, 新, 摆 Europe, student, war, female, wolf, New Zealand, country, prime minister, new, pendulum	0.01
Merkel & Germany	德国, 默, 瑞典, 克, 之声, 政府, 柏林, 中, 书, 出 Germany, Merkel, Sweden, gram, voice, government, Berlin, Chinese, book, out	0.01
India-China disputes	印度, 冲突, 中印, 蒙古, 印, 边境, 边界, 士兵, 军, 量 India, conflict, China-India, Mongolia, India, border, border, soldier, army, quantity	0.01
US-China disputes (1)	驻, 外交部, 发言, 中方, 大使, 捷克, 室, 王, 使馆, 美方 China, Ministry of Foreign Affairs, speech, Chinese, Ambassador, Czech, office, King, Embassy, US	0.01
	知道, 新加坡, 标准, 一个, 不同, 检测, 需要, 承担, 后果, 次	0.01

Singapore	know, Singapore, standard, one, different, test, need, bear, consequence, time	
Wang Liqiang	韩国, 朝鲜, 瑜, 韩, 都在, 一个, 王立, 强, 百分, 心 Korea, North Korea, Yu, Korea, all in, one, Wang Liqiang, percent, heart	0.01
Ant Group	钱, 中, 上市, 蚂蚁, 中国, 诈骗, 局, 一名, 大学, 调查 money, China, listing, Ant, China, fraud, bureau, one, university, investigation	0.01
Masks	口罩, 防疫, 戴, 建议, 瑞士, 越, 措施, 指南, 防护, 民众 masks, epidemic prevention, wear, recommendations, Switzerland, Vietnam, measures, guidelines, protection, people	0.01
Hong Kong protests	香港, 革命, 感谢, 免费, 中, 抗议, 艺术, 百, 时代, 支持 Hong Kong, revolution, thanks, free, China, protest, art, hundred, times, support	0.01
Apple Daily Taiwan	蘋果, 新聞, 網, 台灣, 乘客, 冠狀, 轮, 回, 邮, 公主 Apple, news, web, Taiwan, passenger, coronavirus, round, back, post, princess	0.01
Poland	控, 监, 动物, 恶, 杀, 波兰, 语, 野生, 盛, 监狱 control, prison, animal, evil, kill, Polish, language, wild, Sheng, prison	0.01
Hu Xijin	进, 胡, 梦, 米, 锡, 一直, 中国, 没有, 酒, 世界 Hu Xijin, dream, rice, always, China, no, wine, world	0.01
US-China trade	重, 启, 马, 审, 头条, 主, 判, 聚, 六度, 中美 Re, horse, trial, headline, main, judgment, convergence, six degrees, Sino-US	0.01
Tibet	保护, 宗教, 爱国, 西藏, 信仰, 自由, 文化, 尊重, 续, 意识 protection, religion, patriotic, Tibet, faith, freedom, culture, respect, continued, consciousness	0.01
Mexico	级, 毒, 墨西哥, 解决, 墨, 菲, 注意, 问题, 广播, 锅 grade, poison, Mexico, solve, Mexico, Philippines, attention, problem, broadcast, pot	0.01

References

China Data Lab. 2020. "Baidu Mobility Data." *Harvard Dataverse*.

Hu, Tao, Weihe Wendy Guan, Xinyan Zhu, Yuanzheng Shao, Lingbo Liu, Jing Du, Hongqiang Liu, Huan Zhou, Jialei Wang, Bing She, Luyao Zhang, Zhibin Li, Peixiao Wang, Yicheng Tang, Ruizhi Hou, Yun Li, Dexuan Sha, Yifan Yang, Ben Lewis, Devika Kakkar, and Shuming Bao. 2020. "Building an Open Resources Repository for COVID-19 Research." *Data and Information Management* 4(3): 130–147.

Kinetz, Erika. 2021. "Army of fake fans boosts China's messaging on Twitter." *Associated Press*.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeffrey Dean. 2013. "Distributed representations of words and phrases and their compositionality." *NIPS*: 3111–3119.

Mozur, Paul. 2018. "China Presses Its Internet Censorship Efforts Across the Globe." *The New York Times*.

Mozur, Paul. 2019. "Twitter Users in China Face Detention and Threats in New Beijing Crackdown." *The New York Times*.

Nazeri, Nima, and Collin Anderson. 2013. "Citation Filtered: Iran's Censorship of Wikipedia." *Iran Media Program*.

Roberts, Margaret, Brandon Stewart, and Dustin Tingley. 2016. "stm: R Package for Structural Topic Models." *Journal of Statistical Software*, Publisher: R package version 0.6.

Yamada, Ikuya, Akari Asai, Jin Sakuma, Hiroyuki Shindo, Hideaki Takeda, Yoshiyasu Takefuji, and Yuji Matsumoto. 2020. "Wikipedia2Vec: An Efficient Toolkit for Learning and Visualizing the Embeddings of Words and Entities from Wikipedia." *EMNLP*: 23–30.