Media Coverage and the Success of COVID-19 Response in Democracy

Udi Sommer
And
Or Rappel-Kroyzer
udi.sommer@gmail.com

This work is part of the **PANDEMICS AND POLITICS PROJECT (PPP)**
Forthcoming in *Journal of Information Technology & Politics*

The Center for Combating Pandemics, Tel Aviv University
Center for the Study of the United States at Tel Aviv University with the Fulbright Program
Political Science, History, Tel Aviv University

We examine how internet media outlets in key Anglo-American democracies differed under a similar external shock: the outbreak of the COVID-19 pandemic in early 2020. COVID-19 posed a special challenge to democracy, juxtaposing it with alternative forms of government, which may be better positioned to deal with such a crisis. The online media, as the watchdog of democracy, played a key role. As the pandemic started to spread worldwide, three democracies—the USA, Canada and New Zealand— were of particular interest. The USA had the highest number of cases and deaths, considerably more than its neighbor to the north. NZ was the democracy that most effectively dealt with the pandemic. We comprehensively study the coverage of the outbreak on the internet website of a newspaper of record in each. Data were harvested for the universe of 27,089 articles published online between mid-February and early May on the websites of the *New York Times*, *New Zealand Herald* and the *Globe and Mail*. Natural learning processing and dependency parsing are the methods used to analyze the data. We find meaningful differences between the outlets in timing, structure and content. Compared with their US counterpart, the online watchdogs of democracy in Canada and NZ—where COVID-19 politics were far more effective—barked louder, clearer and two weeks earlier.

Keywords: COVID-19; natural language processing; NLP; pandemic prevention policy; watchdog of democracy; dependency parsing; datamining; timing of media coverage

Introduction

The tumultuous year of 2020 and the COVID-19 pandemic it brought in its wings seemed to pose a key challenge to democracy, juxtaposing it with alternative forms of government. As the disease started to spread globally, questions about the ability of different regimes to deal with its implications started to dominate. As a nondemocracy, the argument went, to withstand the storm China was able to take measures that were hard—or even impossible—for democratic nations to employ. Yet, eventually, contact tracing, quarantine, limitations on movement, infringement on privacy, and a host of other measures were employed worldwide, in democracies and nondemocracies alike. Whereas many democracies used far reaching measures, some of them had more to show for than others in terms of how successful their COVID-19 response was. As the watchdog of democracy, the press and its online outlets were crucially important in general and particularly for the politicization of the crisis (Lilleker et al. 2021; Aelst and Blumler 2021).

The press plays a key role during a political emergency, when pressure runs high and the need for information becomes acute in order to form effective policy. During extreme crises, the role of the press becomes even more salient. Most recently, this was the case with the Coronavirus pandemic. From the very beginning, the online media loomed large. It is hard to overestimate the importance of this democratic institution, which played a crucial role as the interlocutor of COVID-19 information, critical for effective pandemic response. For decisionmakers and the public alike, internet press coverage was a principal source of information.

Our goal is to provide a rich comparative analysis of how key internet news outlets in several Anglo-American democracies differed under a similar external shock: the outbreak of the COVID-19 pandemic. Using cutting edge artificial intelligence algorithms and datamining of

complete data for a period of over 2 months, we examine the role of a key media outlet in the USA, Canada and New Zealand in this political global crisis.

We offer three important innovations for the study of internet news coverage as a crucial player in the political sphere during crisis. First, we describe meaningful differences in how online news outlets in each country operated in terms of timing, structure and content of journalistic coverage. In the case of COVID-19, such differences parallel how well those different democracies coped with the pandemic in terms of producing effective preventative policy. Second, methodologically, we introduce a suite of artificial intelligence algorithms and datamining procedures that make important contributions to the study of politics and media coverage on the internet. Third is the specific external shock we study, the COVID-19 crisis. There is insufficient scholarship on this topic and how it was covered in online media outlets and corresponding policies.

We compare media coverage where the most successful policy response to the COVID-19 outbreak in an Anglo-American democracy was registered, in New Zealand (NZ), to media coverage in the least successful one, the USA. With roughly a quarter of world cases in the USA at that point in the pandemic, the American record is the worst among democracies. Currently, the share in US population is 4012 cases and 80 deaths per 100K. While probably also benefitting from geographical advantages, in NZ, the share of cases in the population is 42 cases and <1 death per 100K. As for Canada, comparable to the USA in terms of its geography, demography, population, economy, trade, education, federal system, natural resources, travel patterns and more—the Canadian policy during to the outbreak of COVID-19 was far better. Currently, share in population in Canada is 984 cases and 32 deaths per 100K.

A range of analyses based on artificial intelligence tools suggest meaningful differences between the internet media outlets in those three democracies in terms of timing, structure and content. Timing is measured in the preemptive capacity of the press concerning pandemic response and the implementation of policy measures. Structure pertains to a range of features of the networks of COVID-19 articles. Content measures the level of politicization in press coverage of the contagion (i.e., policy vs. politics, as we further elaborate below). Compared with their US counterparts, the news outlets in Canada and NZ, where COVID-19 response was far better, were louder, clearer and reacted 2 weeks earlier. While we do not purport to argue for causal relations between internet media coverage and the way the different governments responded to the pandemic, those patterns and correlations are informative and significant theoretically.

Online Media Coverage in times of Crises

The role of the media as the fourth estate—and the crucial role it plays for democracy—becomes clear in light of the decline in some of the other key institutions of democracy. This is particularly evident in times of crisis. Due to a range of factors running the gamut from institutional design and historical legacies to economic factors and the personalization of politics, legislative branches around the world have been in decline for several decades (Ishiyama 2019; Khmelko et al. 2019). The decline of legislatures is accompanied by the concentration of power in executives. Indeed, one of the reasons for the usurpation of legislative power is executive creep at various levels of government (Cooper 2017; Heffernan 2003; Thomas and Lewis 2019). Since human online activity increased even more during the pandemic, cybersecurity risks grew as well (Radanliev et al. 2020). The risks associated with weakening of legislatures are severe

and include economic ramifications (Pelizzo and Baris 2015), and may even amount to undermining the democratic system itself (Khmelko & Wise 2018; Fish 2006). Such risks are more acute in times of crisis, when power grabs are likely (Economist 2020). This is when the crucial role of the press may come to light the most.

Analyses in the scholarly literature of watchdog journalism, however, are less concerned with such circumstances, and apart from a few exceptions (Kim et al. 2020; Lilleker et al. 2021; Aelst and Blumler 2021), they have not yet touched seriously on the case of COVID-19. Yet, in other areas, similar research to ours using datamining and concerning COVID-19, looked at coverage of research on COVID-19 mortality, immunity, and vaccine development by organizations and countries (Radanliev, De Roure, and Walton, 2020). And so, standard accounts of watchdog journalism largely boil down to accountability stories, also known as deep accountability reporting (Stockmann and Gallagher 2011; Marquez-Ramirez et al. 2020). It is the Fourth Estate that is expected to hold representatives to their word and exercise oversight over public officials, who may abuse their power and waste public funds, or simply show poor judgment and flawed decisionmaking. The press is expected to expose scandals, corruption and criminality in the public and private sectors, which in turn would mobilize the public to influence those in power, either directly in the case of elected officials, or indirectly in the case of bureaucrats (Donsbach 1995; Schultz 1998; Norris 2010). Likewise, journalistic coverage may spur corporate as well as legal accountability (Lindstedt and Naurin 2010).

Political affiliations and partisanship also influence the scope and nature of watchdog journalism (Puglisi and Snyder, 2011). Despite economic pressures that come in various forms (Picard 2010; Barnhurst 2011) and challenges from social media, compared to more authoritarian regimes, it is in developed democracies that journalism has been able to keep biting (Stockmann

and Gallagher 2011; Tong and Sparks 2009; Bandurski and Hala 2010). Beyond elected officials, it would also pertain to other political processes, such as constitution making in the European Union (Trenz et al. 2009).

Much of the literature on watchdog journalism is focused on the role perceptions of journalists (Weaver 2007; Beam et al. 2009; Mellado 2015). Those perceptions stem from the professional culture of journalism that include elements pertaining to the independence of the press and the notion that journalism is one form of public service (Kovach and Rosenstiel 2001; Deuze 2002). Prototypical role perceptions for journalists include the detached watchdog, which is most common in established democracies with an independent press (Hanitzsch et al. 2011), as well as the populist disseminator, critical change agent and the opportunist facilitator. Role perceptions, and how journalists perform their capacities, are related to the type of political system in which they operate (Deuze 2002). This perception of the watchdog function of journalism is largely shared by the public (Pew 2011). It is in Anglo-American democracies that the dominant role perceptions of journalists are as watchdogs of the democratic system (Marquez-Ramirez et al. 2020). This kind of democracies are as the focus for our study.

Scholarship has looked at several variables to compare media-politics-policy relations in different countries. Those typically do not include the watchdog capacities of the media as such, but instead look at issues that are either tangentially related or not related at all, such as media markets, links between the press and political parties, the development of the profession of journalism, state intervention and control in the media system and organizations (Blumler and Gurevitch 1995), and media bias (D'Alessio and Allen 2006; Ettema 2007). Some of the scholarship on investigative journalism and muckrakers has taken a micro-level approach to analyzing the function of the media (Clayman et al. 2007). While our approach allows micro-

level analyses as well, the important contribution is the comprehensiveness and depth that datamining and AI allow us in analyzing the interface of policy, politics and online journalism in times of a major 21st Century crisis.

Watchdog journalism has been particularly visible, and in hindsight even glorified (Protess et al. 1991), in times of emergency, such as was the case during the Watergate scandal and the Nixon resignation and during the Clinton impeachment trial (Serrin and Serrin 2002; Shapiro 2003; Berry 2008; Burgh 2008). Our project takes those discussions to the present. We examine how online media operates in such capacities in the Information Age in the midst of a global crisis. To that end, we propose a framework that focuses on timing, structure and content.

Timing, Structure and Content of Online Coverage of the Outbreak of COVID-19

Both in terms of theoretical framework and in terms of empirical reality, the Information Age allows the press to exercise its watchdog functions elaborately in at least three important ways – the timing of news coverage, its content and the structure of its network. This, however, requires rethinking the focus on accountability as the core function of online media.

The focus of extant frameworks on accountability is limited in terms of the potential ramifications for a well-functioning online media. If the outcome is limited to the extent to which the public translates information relayed by the media into punishment for government officials, then the evidence for this is mixed (Pande 2011; Chang et al. 2010). What is more, during an emergency, when time is of the essence, we should be able to examine additional functions of the media.

COVID-19 is a case in point. In such circumstances, the effect of online media should be swift and policy-oriented. Here, instead of reduction in corruption (Chowddhury 2004) or the

replacement of public officials, we would look for improved policy measures, which should be put in place in response to a state of emergency and in a timely fashion. Such policy measures should slow the spread of the pandemic and reduce the number of deaths. Along those lines, the COVID-19 crisis is one where the cooperation of the public is indispensable. Therefore, coverage, for instance, that outlines ways to prepare for certain upcoming policy measures such as quarantine, travel bans, mask wearing and social distancing, is necessary. It will certainly benefit the public and the democratic system as a whole. Yet, it falls outside of the standard accountability framework. An accountability process, by which biting press coverage leads to public concern, which in turn spurs shifts among decisionmakers, may be too slow to fight a global epidemic.

While some list service to the public interest as one of the goals of journalism (e.g., Norris 2014, p. 538), our work adds meaning to this general term. Specifically, we look for timely coverage of policy put in place to help assuage the crisis. It would be hard to demonstrate how press coverage leads to the implementation of any particular policy measure, and in particular in times of emergency. Yet, any correlation between media coverage and government action would be of interest. The role of journalism in informing publics and leaders has ballooned to massive proportions during COVID-19, and certainly in its first few months. Thus, COVID-19 is an important case to develop our understanding of the role of internet media outlets in times of crisis in the Information Age. Let us now delve into those three key elements: Timing, content and structure of media coverage.

Timing – Preemptive Online Journalism

By its very nature, the accountability framework is mostly retrospective in nature. The press holds officials accountable for past actions. Yet, with the amount of information made available

thanks to technological innovations, *prospective* press coverage is not only possible, but may be exercised with ostensibly important consequences. This is the preemptive version of watchdog journalism. In some cases—and in particular in times of emergency and social duress—for the press to play such a preemptive role would be particularly effective.

During January and early February 2020, the sense outside China was that COVID-19 would only indirectly affect the rest of the world. While chains of production may be hampered because of the spread of the disease in Chinese mainland and the lockdowns and quarantines ordered by the government there, it was believed that just like in the case of SARS, the West is going to be spared the brunt of the pandemic. Yet, as cases began to appear outside China—and in particular in countries in the West—the imminence of the threat became clear. Governments that were late to identify the threat and put policy measures in place, often suffered a substantially higher toll (Pei et al. 2020; Barak et al. 2021).

It was the role of the media to pull the alarm, and consequently potentially lead to more effective policymaking. Journalism that can identify a threat and *preemptively* cover the story within the framework of its local and national implications (even when this story is still unfolding overseas) would exercise effective oversight. The pivotal questions here are whether online coverage of COVID-19 expanded before or after the pandemic hit. And, to what extent, and until what point, was press coverage focused on China and world affairs, and at what point did the switch to local and national implications take place?

We expect a more biting press to realize the imminent domestic threat earlier. Since we expect press coverage to correspond with policy measures put in place and with policy outcomes, timing of COVID-19 press coverage in NZ and Canada should be better than in the USA. Accordingly, we would first expect to see expansion in the coverage of the pandemic earlier in

NZ and in Canada. Second, we would expect the press in those countries to put the spotlight on potential domestic implications, as opposed to a more generalized global coverage of the topic, earlier. Lastly, we should expect to see coverage of specific policy measures when this coverage is still relevant in the sense that it could still have implications for the spread of the disease, before it is too late. In sum, in terms of volume of COVID-19 coverage, its focus (local vs. global) and its policy implications, we would expect differences in timing between the *New York Times (NYT)*, *Globe & Mail (G&M)* and *New Zealand Herald (NZH)*.

Structure – Networks and Supernodes

A biting watchdog used to amount to headlines on the front page, in some cases. In other cases, a series of articles, all or most appearing as headlines on the front page. There would also be distinctions drawn between articles that appear above and below the fold and various other structural features. Yet, the structure of internet media outlets is considerably more complex. The way it is used for political purposes is affected accordingly. In particular, like many other areas of human activity, journalism is increasingly organized in the form of a network (Segev 2020). News outlets, and in particular what used to be print journalism, are published in an online format. On the newspaper's website, articles are linked to each other. This suggests several critical ways in which the structure of online media is a marked departure from their old form. The half-life of an article that appears at the top of the website may be pretty short. It is shorter than the 8-12 hours elapsing between the morning and evening editions of the printed format. What is more, the features of the network itself highlight the importance of certain pieces. For instance, while different newspapers exhibit different methods to do that, many of them link articles to each other. In some newspaper websites, links are embedded throughout the articles. In others, there is a dedicated section on the webpage where related articles are linked. This

means that the network within which the article is nested is critical for the traction it would get and for its visibility. The structure of the network is informative and has political consequences.

Using a force-directed layout to model the structure of the network of online media coverage, we can simulate a physical system, whereby the algorithm iterates until it reaches minimal energy when attributing specified attraction forces between linked nodes, on the one hand. On the other hand, the algorithm can specify repulsion forces between each two nodes, both depending on their distance. We would expect the structure of our network to be such that the network behaves much like a social network in which there are relatively few supernodes (articles linked to many other articles) and a lot of leaves (articles linking to one other article). We further explain the Force-Atlas2 algorithm, aimed to simulate the structure of COVID-19 articles, below. But it is designed as an energy model by trying to bring poorly connected nodes closer to very connected ones. Specifically:

Let $G = \{V, E\}$ be a network graph, and let $d: V^2 \longrightarrow \mathbb{R}^+$ be the distance between each pair of nodes.

Then the attraction force $F_a: V^2 \longrightarrow \mathbb{R}^-$ is defined as:

$$\forall e = (v_1, v_2) \in E, F_a(v_1, v_2) = -d(v_1, v_2),$$

and the repulsion force $F_r: V^2 \longrightarrow \mathbb{R}^+$ is defined as:

$$\forall (v_1, v_2) \in V^2, F_r(v_1, v_2) = k_r \frac{(deg(v1) + 1)(deg(v2) + 1)}{d(v_1, v_2)}$$

Modeling the press as a network has been done in the literature. Yet, the methodological innovations introduced here allow us to examine the structure of the press from a range of new angles, including its supernodes, metatags and the almost complete information we harvest about the network and its myriad components (tags, links, nodes, sections etc.). In the network, the nodes and links are meaningful both qualitatively and quantitatively. Qualitatively, we would be

interested in the content of the articles to which a piece is linked (more on that in the Content subsection below). Quantitatively, it is the articles with the most links, the supernodes, which carry more weight in the network. Readers of any article on the topic are probabilistically more likely to click a link that would lead them to a supernode. Those supernodes are interesting since editors' decisions as well as automatic algorithms determine the structure of the network. Thus, the supernodes and the structure of the network as a whole are informative of the political preferences of the elites running the newspapers as well as for the political implications of the press coverage they produce.

The links between the articles are not automatically produced by the papers' CMS systems using a textual similarity algorithm. Rather, there is a significant audience flow through these links since they are topical and substantive in nature. As such, the structure of the network we are able to track is strongly connected to the actual traction of the article. An article is a supernode because it has more purposive links, which is a reflection of the editorial teams' views about what constitutes key work, in addition to the work of automatic algorithms.

We expect to see meaningful differences in the structure of the network and its supernodes in the *NYT*, compared to the outlets in Canada and NZ. As discussed above, COVID-19 was a crisis where informing the public about health and policy measures was of utmost importance. In many ways, this was the key goal of the watchdog of democracy. We would expect the supernodes in the *NZH* and *G&M*, thus, to be more policy-oriented, compared to *NYT*, where we would expect to find more politically-infused supernodes. Along the same lines, we would expect the close networks of supernodes in the *NYT* to be more focused on politics.

Content – Politics vs. Policy

The content of media coverage is critical. This is true for the issues covered, the identity of the people mentioned and the different angles from which those are examined, the valence of the statements relayed and the sentiments expressed. Political aspects are particularly important.

We examine several types of content. First, we look at the extent to which the coverage of COVID-19 was political in nature. In some democracies, and certainly in the USA, many facets of the pandemic turned into political issues. Both in statements by President Trump as well as in reactions of politicians and members of the public, questions related to quarantines, masks and travel bans, for example, were examined from a political viewpoint. Alternatively, under different circumstances or in a different political climate, quarantines, masks and travel bans could simply be a matter of policy, largely driven by health and economic considerations.

When looking at persons and political topics, it would be interesting to see to what extent outlets are focused on political aspects. As further explained below, we are able to conduct a range of AI grammatical analyses that provide a picture of the topics of each sentence, for the topics that dominate an article and for a topic that is likely the focal point for a particular piece. For our purposes, it is interesting to examine the extent to which international politics and political persons are covered as key topics in each outlet.

Materials and Methods

The political effects of online media have been tested using experimental methods (Pande 2011; Chang et al. 2010), econometric techniques using cross-sectional as well as time-series models (Brunetti and Weder 2003; Stapenhurst 2000; Chowddhury 2004; Keefer 2007; Bandyopadhyay 2006; Norris 2012), and case study methodology where the effects of media are studied in specific countries around the world (Jacobs 2002; Reporters without Borders 2007; Francken et

al. 2005; Ferraz and Finan 2008). To the best of our knowledge, ours is one of the first to take advantage of datamining and AI.

We focus on a newspaper of record in each of the countries studied. We scraped all articles from the *New York Times* (*NYT*), The Canadian *Globe and Mail* (based in Toronto) (*G&M*) and *The New Zealand Herald* (*NZH*) websites using selenium, a python scraping framework. For each article, we scraped HTML meta tags, title, author, body, links, images and image captions. In the *NYT* we then filtered only articles with the meta type PT(piece type)=article, to filter out videos and interactive material which is not usable and can distort our analysis (such as detailed election results). The *NZH* and G&M websites are simpler, so this filter was unnecessary. To determine which articles are relevant to COVID-19, the "news keywords" metatag in the article's HTML was used. Using this tag, the website owner or editor are able to inform news aggregators, particularly Google news, under what keywords the article should be grouped. While Google has officially dropped support for this meta tag, it is still widely used by websites, specifically the ones we researched. This method is superior to searches over the article's title or body, as it follows the decision of the site editor herself.

This framework provides a considerably more complete picture of the network. We do not look at specific articles assuming they represent the entire corpus, nor do we even look for all the data concerning the subject of discussion (as done, for instance, by *Lexis-Nexis*). Rather, we extract the universe of newspaper articles. This allows us to perform analyses on the share of texts concerning our subject of the entire corpus, and compare between texts that concern our subject and those that do not. There is no sampling involved, and thus biases typically associated with inference are avoided. Furthermore, as we download and scrape the entire archive, the risk of having a biased selection of texts to analyze is marginal. While obviously the web archive

does not contain everything ever published, texts which are missing or have less informative metadata are either older or in places the website owner put less emphasis on. We avoid those articles that the elites involved (e.g., editors and owners, politicians, regulators, campaigners) were least interested or vested in. Lastly, and relatedly, by avoiding a third-party tool (such as *Lexis-Nexis*) we have knowledge and control over what was retrieved, how and why. There is less risk of selection biases, which the researcher is not aware of, that would be embedded within the third-party tool. In other words, since we control the process, it is transparent to our readers as well as to us. We are not dependent on algorithmic decisions of any other entity providing the data.

From February to May, the harvested *NYT* pieces totaled 11,213, of which 4,917 were COVID-19 articles. The 5 most popular sections included US (1768 articles), Opinion (1283), World (946), Arts (905) and Business (817). On weekdays, the NYT averaged 176 articles a day, and on weekends 77. On the *NZH* website, we harvested 8,404 articles, of which 2,951 were COVID-19 articles. The 5 most popular sections included NZ (3146 articles), Business (1344), Sport (967), World (909) and Lifestyle (727). The *NZH* weekday daily average was 123 articles and 78 on weekends. On the *G&M* website, we harvested 7,472 articles, of which 3,382 were COVID-19 articles. The 5 most popular sections included Canada (2205 articles), World (1369), Sports (855), Business (722) and Opinion (526). The *G&M* weekday daily average was 122 articles and 43 on weekends. On top of the standard textual fields such as article title, subtitle and full text, our scraping framework enabled harvesting of metadata invisible to the user but highly informative of the article, such as its tags, how it is summarized for social media entries, etc. On top of this information, we harvested image captions and links. The latter enabled constructing the structure of the network.

We deploy natural language processing (NLP) algorithms to enable deeper analysis of the texts, the article title, description, body and more. We used the SpaCy[1] python library, in which the basic language pipeline consists of the following steps:

1. Tokenizing: splitting the text into tokens – words, punctuation etc. This step is performed using a rule-based algorithm, based on splitting the text by spaces and then applying language specific rules to prefixes, suffixes and whether periods are used as separators.

2. Part-of-speech tagging: assigning to each token its part of speech tag: noun, verb, pronoun etc. This is done using a perceptron algorithm—a simple supervised learning algorithm—which uses a SpaCy's English Convolutional Neural Networks (CNN)—a statistical model trained on millions of web pages in English (Koo and Collins 2008, Honnibal 2013). For highest accuracy, we used the largest model provided by SpaCy[2]. This is the default model delivered in Spacy and it perfectly fits our needs as it was trained on web pages, which is the nature of our data as well; it is used in all of the following steps as well.

3. Dependency parsing: This is the most complex stage of the process, whereby the syntactic structure of the sentence is understood, and a parse tree is created. The root is the sentence verb, followed by the tokens dependent on the root, then each followed by the tokens dependent on it and so forth. The parsing is done using an algorithm that combines greedy transition-based parsing and CNNs (Honnibal and Montani 2017).

The parse tree in Diagram 1 is an illustration for the final product at the end of this stage from an actual March 13, 2020 title in the *NYT*.

---

[1] SpaCy 2.2.4, https://spacy.io/
[2] https://spacy.io/models/en

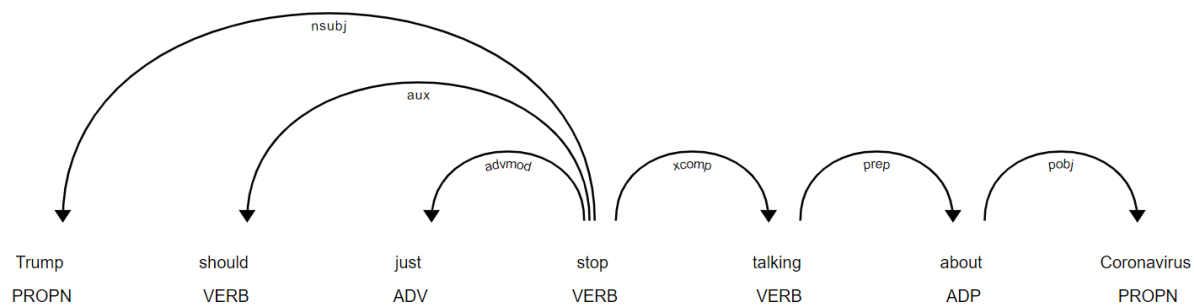| Trump	should	just	stop	talking	about	Coronavirus
PROPN	VERB	ADV	VERB	VERB	ADP	PROPN

Diagram 1: Example of Parse Tree (NYT, March 13 2020)

Furthermore, we also apply the next step in the NLP pipeline, named entity recognition:

4. Entity recognition finds tokens, which correspond to real-world entities, such as persons, geopolitical units, books, laws, etc. It uses an incremental parsing with bloom embeddings and residual CNNs, in a similar approach to the dependency parsing step.

The results of this step however vary in quality, and thus we used only the Person results. On the data we scraped or extracted through AI we have performed various histograms and used the Agresti-Coull algorithm (Agresti and Coull 1998), implemented in the statsmodels python package (Skipper and Perktold 2010) to determine confidence intervals. The network structure of the articles and links were visualized using the Force Atlas algorithm in Gephi (Jacomy et el. 2014).

Results

Timing - Preemptive Online Journalism

In the case of COVID-19, for a media outlet to be preemptive, it would cover the pandemic within the framework of its national implications, even when the story was still unfolding overseas. This kind of coverage, switching from international politics to domestic politics, may

have critical policy implications. How dominated the press has been with COVID-19, and how quickly that gathered momentum, is clear from Figure 1. The lines with the shaded margins of error depict the share of COVID-19 articles—determined by the "news keywords" meta tag—out of the overall daily number of articles. *NYT-NZH* comparison is at the top panel and *NYT-G&M* comparison is at the bottom. The number of total cases is measured on the vertical axis on the right in logarithmic scale and is indicated by solid lines. Patterns over time of the share of COVID-19 articles in each newspaper are quite similar. All outlets offered limited coverage of the pandemic initially (<15% of articles). Yet, by mid-March all quickly expand their coverage. From mid-March until late April, for instance, *NYT* COVID-19 coverage hardly drops below 60%, with multiple days where at least 7 out of 10 articles focus on the pandemic (the dip in mid-April is due to a parenting project ran in the *Times* on those specific days).

Despite those similarities, discrepancies in the timing of the expansion of coverage are particularly informative. *NZH* coverage (blue line) climbed a week *before* the pandemic started spreading in NZ (purple). Likewise, *G&M* coverage (green) preceded the surge in Canada (light green) and coincided approximately with that in the USA. Conversely, increase in *NYT* coverage (red) took place 1-2 weeks *after* COVID-19 started spreading in the USA (orange). The press in NZ and Canada had far better timing.

[Figure 1 here]

During January and early February 2020, the sense in the West was that because of chains of production, COVID-19's effect in the rest of the world would be mostly on business. The main impact was going to be in China, whereas the West was going to be spared the brunt of the pandemic. At what point were news outlets able to entertain the possibility of the pandemic reaching their shores and accordingly consider national implications? Figure 2, which compares

COVID-19 coverage in different sections of the newspapers, demonstrates substantial disparities. The most popular sections are included in the figure. Red lines mark domestic news sections; Yellow, green, and blue mark the World, Business and Opinion sections respectively.

A good baseline is the share of articles in the Business section out of COVID-19 articles, which was remarkably similar at 15-20% in all three outlets. This is in line with the initial expectations that COVID-19 would mostly affect business. Yet, in the period leading up to the outbreak at the national level, compared to their US counterpart, the NZ and Canadian media reacted earlier, not just in the volume of COVID-19 articles (Figure 1), but in content as well. The share of World section articles among *NYT* COVID-19 coverage (top panel in Figure 2) was approximately 50% in mid-February, at a point when the share of US/NYregion articles was less than 7%. It is not until the first week of March that COVID-19 articles in the US section significantly outnumber World articles. On those same weeks, in non-COVID-19 *NYT* articles, the World section was 7-9% and the US/NYregion sections were 17-23%.

Contrariwise, COVID-19 coverage in *NZH* (bottom left panel) and *G&M* (bottom right) touched upon international and national themes equally as early as February. Indeed, closer scrutiny we conducted of article content suggests that several of the *NZH* pieces in this early period focused on potential implications if COVID-19 reached NZ. Likewise, and related to the structure of the network of *NZH* COVID-19 articles discussed later, all top 15 *NZH* COVID-19 supernodes appeared in the NZ section. This suggests that not only timing was oriented toward domestic implications, but the same was true for the structure of the network, as reflected in its supernodes (more on that in the Structure section below).

[Figure 2 here]

Figure 3 shows the share of mentions of "travel ban" and synonyms (such as flight ban) and "quarantine" and synonyms (such as lockdown) in article titles. Coverage of travel bans (left panel) reveals further evidence for the importance of timing and the preemptive nature of press coverage. Travel bans are frequently mentioned in the *NZH* up until March 21, when first NZ cases without direct link to travel abroad are diagnosed. In the *NYT*, there is no mention except for when president Trump announced a travel ban from 26 European countries (March 12). The *G&M* pattern resembles that of the *NYT*. *NZH* discussed travel bans, when it was still an effective policy tool to stem the pandemic in NZ. Motivated, apparently, by political events (Trump's declaration), *NYT* coverage appears late and at a point in time when such policy is rendered ineffective. Quarantine patterns are similar (right panel). *NZH* coverage significantly increases on March 21, when the 4-level alert system was introduced and four days before NZ entered full shutdown. *NYT* title share remains close to 0 over the entire period. The *NYT-G&M* comparison yields no meaningful difference.

[Figure 3 here]

Structure – Network and Supernodes

Figure 4 compares the network of COVID-19 articles in each outlet, with nodes representing articles, and edges representing hyperlinks between them. Node size is proportional to the number of incoming links and node color indicates content on a politics vs. policy scale. The content scale was determined by news_keywords meta tags. Colors depict content of articles (and links) ranging from political (red – 100% of the news_keywords meta tags are political) through politics-policy balanced (yellow – 50%-50%) to those focused on health and economic policy (blue - 100% of the news_keywords meta tags relate to policy). We give several examples for this politics-policy continuum later on. All 4,917 *NYT* (Panel A), 2,951 *G&M* (Panel C) and

20

3382 *NZH* (Panel D) COVID-19 articles from 2/15-5/1 are included. Panel B depicts a subnetwork of a *NYT* supernode.

The *NYT* network (Panel A) is distinct in several ways. First, on the spectrum of topics from political (red) to policy, economy or health (blue), many of the articles are even (yellow) or tilting political. Second, it is very dense. The algorithm simulates a physical system, whereby it iterates until it reaches minimal energy when attributing specified attraction forces between linked nodes and repulsion forces, both depending on distance. The *NYT* website uses inter-article links more often, and their articles of the entire politics-policy spectrum are evenly spread, meaning that political articles get links from all sorts of articles. The Supernodes—which are a key indicator for the focal points of press coverage—are by comparison more political in this outlet. The title of the second most linked *NYT* COVID-19 supernode, "He Could Have Seen What Was Coming: Behind Trump's Failure on the Virus", epitomizes the politically-infused nature of the majority of *NYT* COVID-19 supernodes. The close network of this supernode (Panel B) was heavily political too. Furthermore, only 2 of the 15 *NYT* supernodes deal with clarifications and instructions on how to deal with the pandemic or lockdown. The rest are political.

By comparison, in the *G&M*, one major supernode which is politically neutral (yellow), is titled "How many coronavirus cases are there in Canada, by province, and worldwide? The latest maps and charts". Almost all the rest of the graph is trending blue, and indeed all of the remaining top 14 supernodes. A political figure—Justin Trudeau—appears only in two supernodes. Two other supernodes depict actions by provincial authorities rather than stating names of specific politicians. The remaining 9 supernode articles are informative in nature.

Three of these articles feature Andres Picard, the newspaper's health reporter, answering reader questions.

Lastly, compared to the *NYT*, the *NZH* network (paned D) is sparer, and much more clearly split into a more political cluster and a less political one. In the structure of the network of COVID-19 articles in the *NZH*, political articles tend to link to each-other, and not intermix with other articles. When one read through a non-political COVID-19 article in the *NZH*, there was little chance that they would be directed to a political one. Likewise, the supernodes are less political. Of the top 15 supernodes in the NZH, 8 report on the current state of the pandemic in NZ. Titles of such supernodes include: "Student tests positive, Logan Park High School closes for 48 hours" or "No quarantine for 3600 people returning to NZ from overseas". 4 of the 15 supernodes issue instructions and advice, such as "What Covid-19 alert levels 3 and 4 mean for you and your family" or "Perspex screens and pack-your-own, Countdown's lockdown safety measures". A political figure appears in one supernode title: "Eight new cases, Jacinda Ardern says don't panic, but be prepared". Additional supernode titles discuss frontline health workers' positions and expert accounts.

[Figure 4 here]

Content – Politics vs. Policy

Based on automatic sentence parsing and entity recognition, in Figure 5, we present a word cloud histogram of title subjects, where the colors depict categories of the grammatical subjects: political terms in red; geopolitical entities in purple; and, in blue policy concerning health, science and the economy; grey is for other.

[Figure 5 here]

The term "Coronavirus" has approximately the same share in all the word clouds, and is thus a useful baseline, statistically and substantively. Such a baseline is intuitive and makes sense; Coronavirus appears as the grammatical subject of an article at approximately the same rate in each of the three outlets. However, in other respects, the outlets differ, and meaningfully so. As we expected, *NYT* coverage is heavily political (red font), with 19% of the *NYT* title subjects within the politics realm compared to 14% in *NZH* and *G&M*. The Canadian newspaper of record is focused on policy subjects (blue font) and geostrategic subjects (purple font). The *NZH* is policy-focused, with 51.5% of titles in the policy realm, compared to 37-47% in the other two publications. See also the Online Appendix for additional analyses.

Discrepancies in contents of coverage—and how politicized that coverage may be—is clear from the persons mentioned in COVID-19 pieces. According to the entity extraction algorithm, as we predicted for the *NYT*, of the 25 most commonly mentioned persons in article titles, 20 are politicians – 14 American (e.g., Donald Trump, Andrew Cuomo, and Joe Biden) and 6 foreign (e.g., Boris Johnson, Xi Jinping and Angela Merkel). Only one is a health official – Dr. Anthony Fauci. In article bodies, all but one of the top 15 persons are politicians.

Conversely, in the *NZH*, of the 25 most commonly mentioned persons, only 11 politicians appear in article titles: 8 local (e.g., Jacinda Ardern, Simon Bridges and Grant Robertson) and 3 foreign (Donald Trump, Boris Johnson and Scott Morison). Doctors or public officials appear 5 times more frequently than in *NYT* titles, including Ashely Bloomfield, Dr. Gary Payinda, and Dr. Lance O'Sullivan. In article bodies, 2 more foreign politicians join the list (Andrew Cuomo and Mike Pense) and 4 doctors: Dr. Skegg, Dr. Ghebreyesus, Dr. Mcelnay and Dr. Fauci.

Discrepancies are particularly clear with respect to the Head of Government. In all three cases, it is a charismatic and dominant leader. Ms. Ardern, Mr. Trudeau and Mr. Trump have

captured the attention of elites and the public at home and abroad. Yet, the prominence of each in press coverage is different. Figure 6 compares the share of leaders' appearance as grammatical subject in the title, description (an article component designed to summarize article content for social media), body and image captions. Black whiskers indicate 2 standard deviation confidence intervals. Trump's share in the title and the description of *NYT* articles is significantly higher than that of Ardern in the *NZH* or Trudeau in the *G&M*. The only exception is image titles, where Ardern's share is significantly higher than Trump's, which is true for Trudeau as well. See also Online Appendix.

[Figure 6 here]

Relative to the press in NZ and Canada, the *NYT* is also more engrossed with international politics. In Figure 7 is the share of the terms "China" and "WHO" (World Health Organization) out of the entire number of COVID-19 articles (confidence interval whiskers in black).

[Figure 7 here]

China and the WHO, respectively the foreign country and intergovernmental organization most pertinent to the international politics of the pandemic at the outbreak of COVID-19, are covered in the *NYT* much more heavily than in the *NZH*, as indicated by the share of their appearance in article titles, descriptions and bodies. The *NYT-G&M* contrast is milder.

The COVID-19 pandemic, and the decisive role internet media coverage played in its first few months, make for a compelling analysis of the interface of politics and information technology. We provide a quantitative comparative framework for COVID-19 online coverage (Price 2020; Waisborg 2000; Walgrave, Soroka and Nuytemans 2003). Aided by cutting-edge methodologies, our findings suggest distinct patterns in timing, structure, and content of news coverage. Those differences correspond with COVID-19 policy response in each country.

Compared to the internet media coverage in NZ and Canada, the *New York Times* was relatively late not only to identify the threat, but also to consider it in terms of its domestic implications. In terms of content, the internet websites of the Canadian and New Zealand outlets exhibited greater focus on policy issues, and were less concerned with politics, nationally or internationally.

A caveat to our discussion of the timing dimension relates to global effects. Beyond the national infection rate, the global picture may also be important. It is possible that the online outlets also respond to the global picture, or at least the regional one in the case of the *G&M*, where the effect on Canada's southern neighbor may have influenced coverage in the Canadian media.

Conclusions

Democratic powers, led by the USA and Western Europe, were the big winners of the second half of the 20th Century. With the collapse of the Soviet Union and the end of the Cold War, some argued democracy was the greatest thing that happened in that century. Yet, at the dawn of the 21st Century, the fortunes turned. Starting with the 9/11 attacks on the United States, the resilience of Democracy was tested. The financial crisis of 2007-08 faced democracy with one of its greatest challenges; public trust in this form of government started to wane. The ability of nondemocratic regimes to recuperate from the crisis quicker and more effectively put to question the capacity of democracies to deliver for their citizens. With its system of government, China was better positioned to withstand the crisis and recover at its wake. Next came Russian intervention in the 2016 US presidential elections. The ability of a nondemocratic regime to interfere with an institution as fundamental to democracy as elections—and potentially change

its upshot—again put to question the robustness of this form of government. Cooccurring with the Eurozone crisis and Brexit that same year, democracy's strength was further doubted.

The tumultuous year of 2020 and the COVID-19 pandemic it brought in its wings seemed to pose a similar challenge to democracy, juxtaposing it with alternative forms of government. As the disease started to spread globally, questions about the ability of different regimes to deal with its implications started to dominate. As a nondemocracy, the argument went, to withstand the storm China was able to take measures that were hard—or even impossible—for democratic nations to employ. The free press, as we showed, played a role in democracies, and in their success or failure to deal with the pandemic.

The press is the watchdog of democracy. In the COVID-19 crisis, the media played an expanded role as information about policies, best practices and key events was largely relayed through the press and influenced both public opinion and elite decisionmaking. If we accept agenda setting theory, the media is consequential for *what* we think about, and along the lines of framing theory, *how* we think about it. In a reality infused with closures and limited mobility, online media was particularly influential. Although a media outlet largely critical of the administration, the politicized manner the Trump Administration employed in its crisis management tactics was echoed in much of the coverage provided by the *NYT*. Indeed, while the NYT might have a clear political view, so does any other paper in the US. In terms of the media's preemptive capacity, politicization of coverage, treatment of policy measures and the supernodes in the COVID-19 article networks, coverage in the *NZH* was distinctly different.

The timing, politicization of content and structure of the watchdog of democracy are particularly consequential in times of political crises such as COVID-19. With the caveat of causality still unsubstituted, future work should delve into the relations between politics and

pandemics. And it should look at additional aspects such as partisanship, gender, race and education - all potent politically and in particular in times of pandemics. In that context, it would be useful to further delve into the interface of information technology and politics to learn how our findings may have changed as the pandemic proceeded. This would be particularly interesting with changes in government that took place a few months into the pandemic, at least in the USA. Is it possible that as the Biden administration was ushered in, the nature of the coverage of the pandemic changed in terms of timing, content and structure on online outlets such as the *NYT*? One interesting example here would be to examine the coverage of the flagship policy of the administration in its first year in office, the vaccination campaign.

As for the methodological innovations we propose, the infrastructure of full-population collection and the statistical benefits of sampling merit further and more nuanced discussion in later work. Future work will involve collecting more papers per country to further explore this (Kellam and Stein 2015; Mazzoleni 2010). Additionally, the scrape is closer than *Lexis-Nexis* to the source (the CMS) but it is still temporally dependent. While we may not completely control the process, we get close enough.

A key limitation of the methodologies used is the small N. We study three cases. Furthermore, algorithmic analysis such as NLP, by the nature of things is not foolproof. Yet, this project highlights the benefits of using NLP in the social sciences writ large, and in particular in the study of the interface of information technology and politics. This is doubly true in the context of COVID-19, where digitization of information consumption was enhanced because of social distancing and closures, which lead people to consume their information, political and otherwise, on digital platforms even more. Finally, the analysis of content is based on judgement

calls about the definitions of what accounts as policy versus what is considered political issues, which may also be a downside.

The range of innovations presented here are useful for our understanding of online coverage of political events related to external shocks of unfamiliar nature even beyond COVID-19. As the world prepares for dealing with the pandemic in the long haul, the importance of the media is unlikely to subside in this Information Age. Furthermore, such external shocks may take the form of a climate catastrophe. Our frameworks for the interface of policy and internet during crisis in the Information Age, and the crucial role of the media in this nexus, may have particular analytical usefulness under such circumstances.

References

Agresti, A., and B. A. Coull, "Approximate is better than 'exact' for interval estimation of

  binomial proportions", The American Statistician 52, 119–126 (1998)

Bandurski, D. and Hala, M., eds. 2010. Investigative Journalism in China. Washington:

  University of Washington Press.

Barak, N., Sommer, U., and Mualam, N. 2021. "Urban attributes and the spread of COVID-19:

  The effects of density, compliance and socio-political factors in Israel". Science of the

  Total Environment

Burgh, H. D., ed. 2008. Investigative Journalism. London: Routledge.

Cooper, Christopher A. 2017. "The Rise of Court Government? Testing the Centralization of

  Power Thesis with Longitudinal Data from Canada." Parliamentary Affairs 3 (1): 589–

  610

D'Alessio D. and M. Allen. 2006. "Media Bias in Presidential Elections" *Journal of

  Communication*, 50(4): 133-56

Donsbach, W. 1995. "Lapdogs, Watchdogs and Junkyard Dogs." Media Studies Journal, 9: 17–

  30.

Economist, April 23rd 2020, "A pandemic of power grabs" retrieved on Aug 5 2020,

  <https://www.economist.com/leaders/2020/04/23/autocrats-see-opportunity-in-disaster>

Ettema, J. "Journalism as reason-giving: Deliberative democracy, institutional accountability,

  and the news media's mission", Political Communication, 24, 143-160 (2007)

Ishiyama, J. 2019. "Conclusion: Toward a Theory of Legislative Decline" PS: Political Science

  & Politics, 52(3):277-280

Jacobs, S. 2002. "How Good Is the South African Media for Democracy? Mapping the South
African Public Sphere After Apartheid." African and Asian Studies, 1:279–302.

Jacomy, Mathieu, Venturini, Thomaso, Heymann, Sabastien and Bastian, Mathieu (2014)
ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization
Designed for the Gephi Software. PLoS ONE 9(6): e98679.

Keefer, P. 2007. "Clientelism, Credibility, and the Policy Choices of Young Democracies."
American Journal of Political Science, 51:433–48.

Kellam, Marisa, and Elizabeth Stein. 2015. "Silencing Critics: Why and How Presidents Restrict
Media Freedom in Democracies." Comparative Political Studies 49 (1): 36–77.

Khmelko, Irina and Charles R. Wise. 2019. "Introduction: The Decline in Legislative Powers and
Rise of Authoritarianism", *PS: Politcal Science and Politics*, 52(2): 267

Khmelko, Irina, Rick Stapenhurst, Michael L. Mezey. 2019. Legislative Decline in the 21st
Century: A Comparative Perspective. Routledge

Kim. E. Michael E. Shepherd and Joshua D. Clinton. 2020. "The effect of big-city news on rural
America during the COVID-19 pandemic" PNAS September 8, 2020 117 (36) 22009-
22014; first published August 20, 2020

Knobel, Beth. 2018. The Watchdog Still Barks: How Accountability Reporting Evolved for the
Digital Age (Donald McGannon Communication Research Center's Everett C. Parker
Book Series)

Koo, T., X. Carreras, M. Collins. "Simple semi-supervised dependency parsing." *Procedures of
ACL-08: HLT* (2008).

Lilleker Darren, Ioana A. Coman, Miloš Gregor, Edoardo Novelli. 2021. Political
Communication and COVID-19: Governance and Rhetoric in Times of Crisis. Routledge

Norris, P., ed. 2010. Public Sentinel. Washington, DC: The World Bank.

Pei, S. and S. Kandula, J. Shaman, "Differential Effects of Intervention Timing on COVID-19 Spread in the United States" *medRxiv* (2020)

Radanliev, P., De Roure, D., Walton, R. et al. Artificial intelligence and machine learning in dynamic cyber risk analytics at the edge. *SN Appl. Sci.* 2, 1773 (2020). https://doi.org/10.1007/s42452-020-03559-4

Radanliev, P., D. De Roure, and R. Walton. 2020. "Data mining and analysis of scientific research data records on Covid-19 mortality, immunity, and vaccine development - In the first wave of the Covid-19 pandemic." Diabetes & Metabolic Syndrome: Clinical Research & Reviews 14(5): 1121-32

Schultz, J. 1998. Reviving the Fourth Estate: Democracy, Accountability, and the Media. Cambridge: Cambridge University Press.

Segev, E. (2020). "Textual Network Analysis: Detecting Prevailing Themes and Biases in International News and Social Media". *Sociology Compass*, 14(4), 1–14.

Stapenhurst, R. 2000. The Media's Role in Curbing Corruption. World Bank Institute Working Papers.

Stockmann, D. and Gallagher, M. E. 2011. "Remote Control: How the Media Sustain Authoritarian Rule in China" *Comparative Political Studies*, :1-32

Thomas, Paul E.J. and J.P. Lewis. 2019. "Executive Creep in Canadian Provincial Legislatures" *Canadian Journal of Political Science*, 52(2):363-383

Tong, J. R. and Sparks, C. 2009. "Investigative Journalism in China Today". Journalism Studies

Van Aelst Peter and Jay G. Blumler. 2021. Political Communication in the Time of Coronavirus. Routledge

Weaver, D.H. 2007. The American Journalist in the 21st Century. Mahwah, N.J.: L. Erlbaum

Associates.

Appendices

Online appendix in separate file
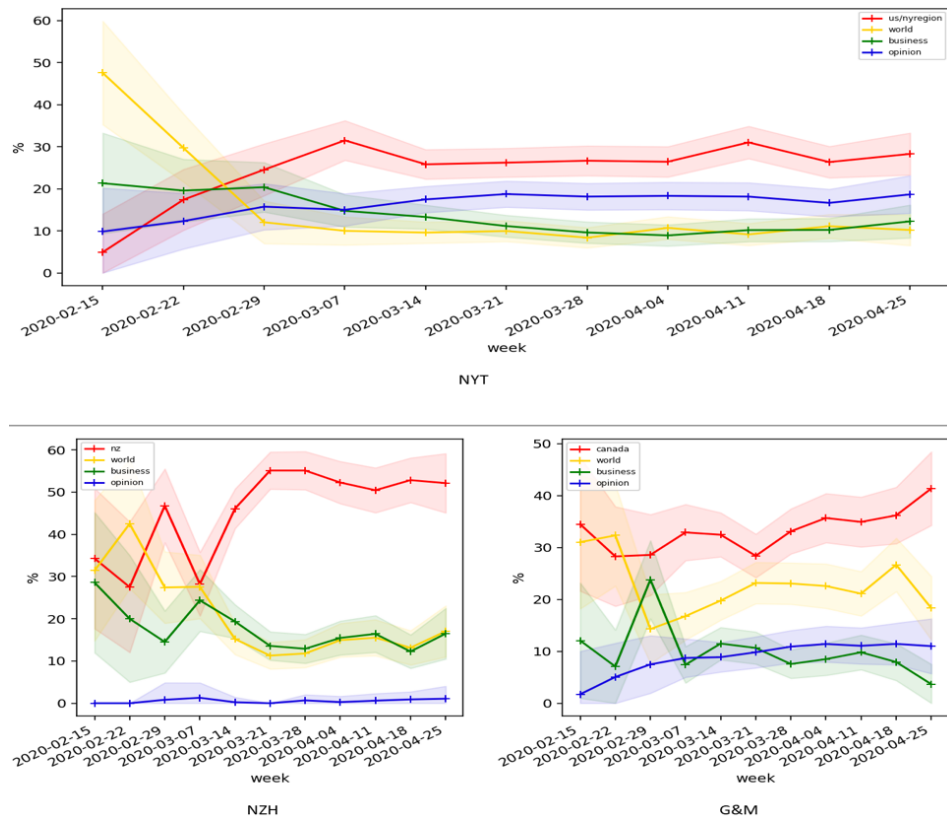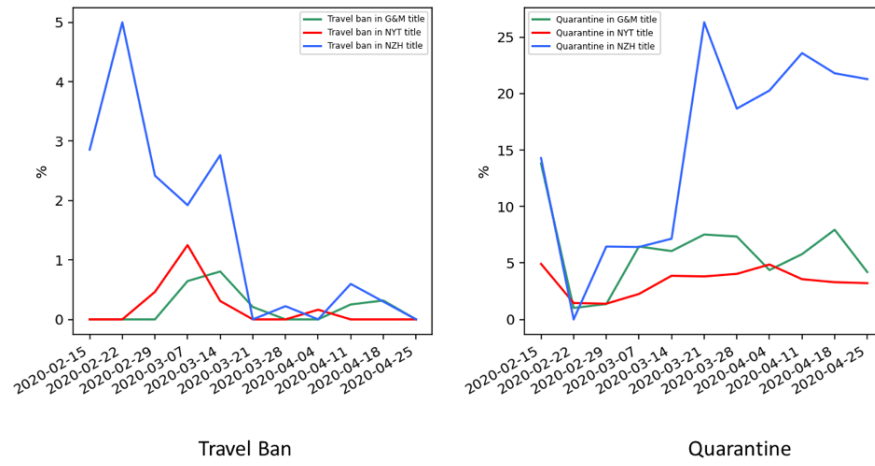
Tables

None

Figures

Figure 1

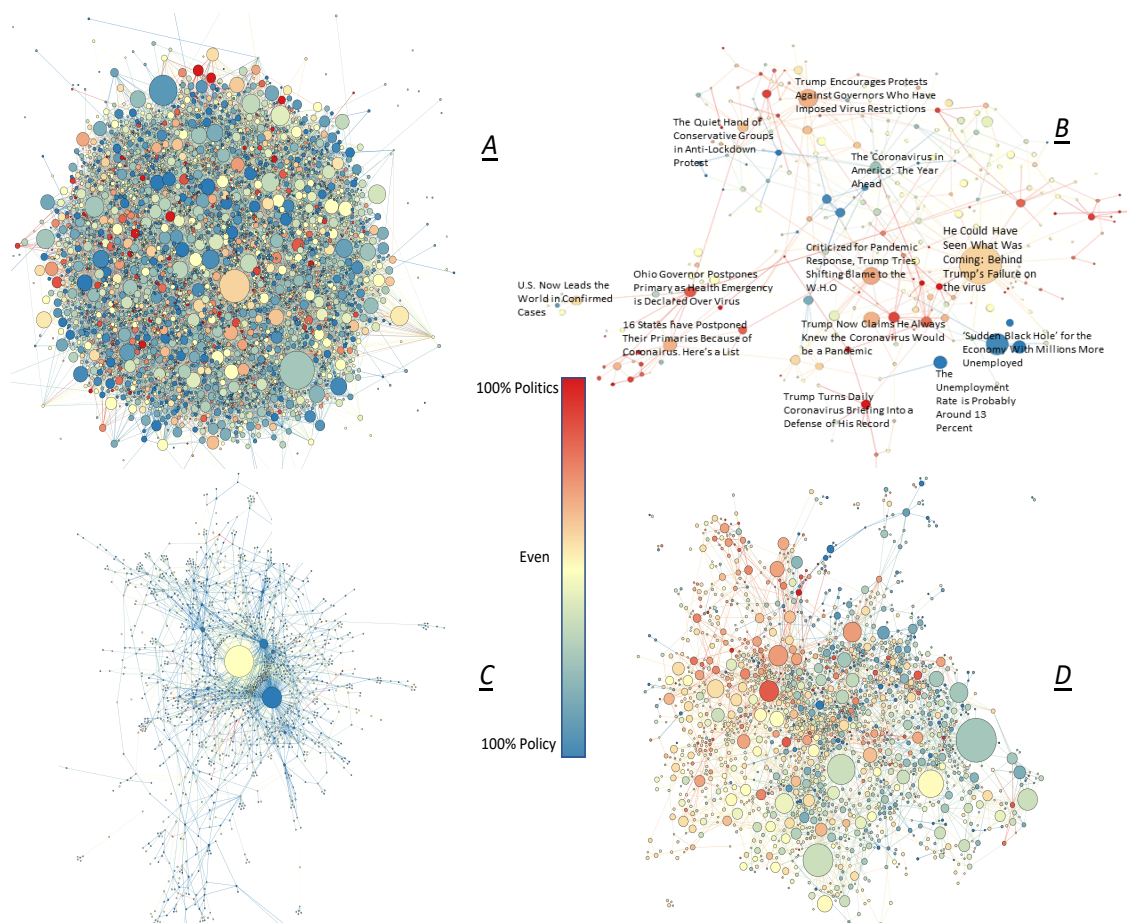NYT vs. NZH

NYT vs. G&M

Figure 2

NYT

NZH

G&M

Figure 3

Travel Ban

Quarantine

Figure 4

Text labels in panel B:

Trump Encourages Protests Against Governors Who Have Imposed Virus Restrictions

The Quiet Hand of Conservative Groups in Anti-Lockdown Protest

The 'Coronavirus in America: The Year Ahead

He Could Have Seen What Was Coming: Behind Trump's Failure on the virus

U.S. Now Leads the World in Confirmed Cases

Ohio Governor Postpones Primary as Health Emergency is Declared Over Virus

Criticized for Pandemic Response, Trump Tries Shifting Blame to the W.H.O

16 States have Postponed Their Primaries Because of Coronavirus. Here's a List

Trump Now Claims He Always Knew the Coronavirus Would be a Pandemic

'Sudden Black Hole' for the Economy With Millions More Unemployed

The Unemployment Rate is Probably Around 13 Percent

Trump Turns Daily Coronavirus Briefing Into a Defense of His Record

Color scale legend:
100% Politics
Even
100% Policy

Panel labels: A, B, C, D

Figure 5

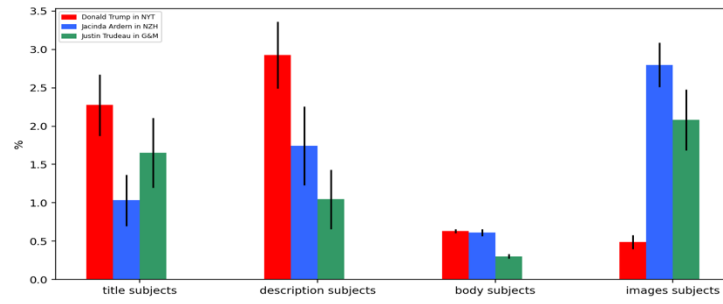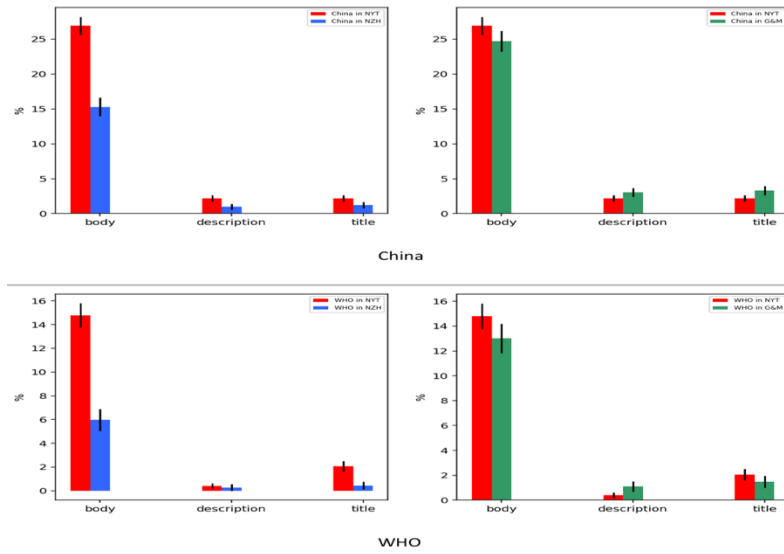**Figure 6**

Figure 7

China

WHO

Figure Captions

Fig. 1: **COVID-19 Number of Cases and Share of Overall Daily Articles**

Fig. 2: **Share of each Section in Weekly COVID-19 Coverage**.

Fig. 3: **Press Coverage of Policy Measures (Quarantine and Travel Ban).**

Fig. 4: Structure of COVID-19 Articles in the Different Internet Outlets.

Fig. 5: Grammatical Subjects in Titles of COVID-19 Articles.

Fig. 6: Heads of Government as Grammatical Subjects in Online COVID-19 Coverage.

**Fig. 7: Coverage of International Politics Topics (WHO & China).**

Appendix

Politicized Coverage
As the results in the article suggest, the *NYT* fundamentally differs from its two counterparts in the extent of politicization of COVID-19 coverage. Coverage in the *NYT* is politically framed, more so than in the *G&M* and with an even more substantial difference from the *NZH*.

Further substantiation to this difference between the outlets is found in terms of the types of grammatical subjects their articles are focused on. A strong indication for the content of the article is when an item is the grammatical subject of a sentence in the body of the article, and is also the grammatical subject of its title. For instance, it makes sense that the head of government would be discussed in the text, but if he or she are not the grammatical subject of the title, the piece is probably not about him or her. For example, President Trump is mentioned in the body of an article titled *Coronavirus Infection Found After Cruise Ship Passengers Disperse*, yet the article is not about the president. But, an article on the same topic titled *Trump Was Furious That Passengers with Coronavirus Were Brought Back to U.S.* is focused on Mr. Trump. Additionally, if a term is rare, we would expect this ratio to be high (if a rare term is mentioned in the body, it is also likely to be at the article's core). Accordingly, we applied this score on all terms which are the grammatical subject of a sentence in at least 10 titles.

Considering coronavirus as the baseline for grammatical subjects, the rate of this term in the *NZH* was 25% (Fig. S1, Bottom-Left Panel) in the G&H ~15% and in the *NYT* ~12% (Fig. S1, Top Panel). The top terms in the *NYT* are overwhelmingly political or geopolitical entities (i.e., a nation-state or a state, a political figure or a political institution or office). Of those political and geopolitical terms, Europe, the USA, the UK and the Federal Reserve, all appear even more frequently than the Coronavirus baseline. With scientific and Policy issues dominating, the coverage in the *NZH* is remarkably different. Subjects include: Experts, Quarantine, Doctors, Schools and Tolls. Only 4 of the top 15 subjects are political including Council, referring to the various city councils and New Zealand. PM Ardern is only the 11th most popular term. The difference between the *G&M* and *NYT* coverage is less stark, yet the NYT is clearly more political than the G&M, where it is geopolitical entities and policy issues that are most frequently the grammatical subjects. Depicting Title to Body ratio of grammatical subjects in COVID-19 articles, Figure S2 tells a similar story about the relative politicization on *NYT* coverage, compared to *G&M* and *NZH*.
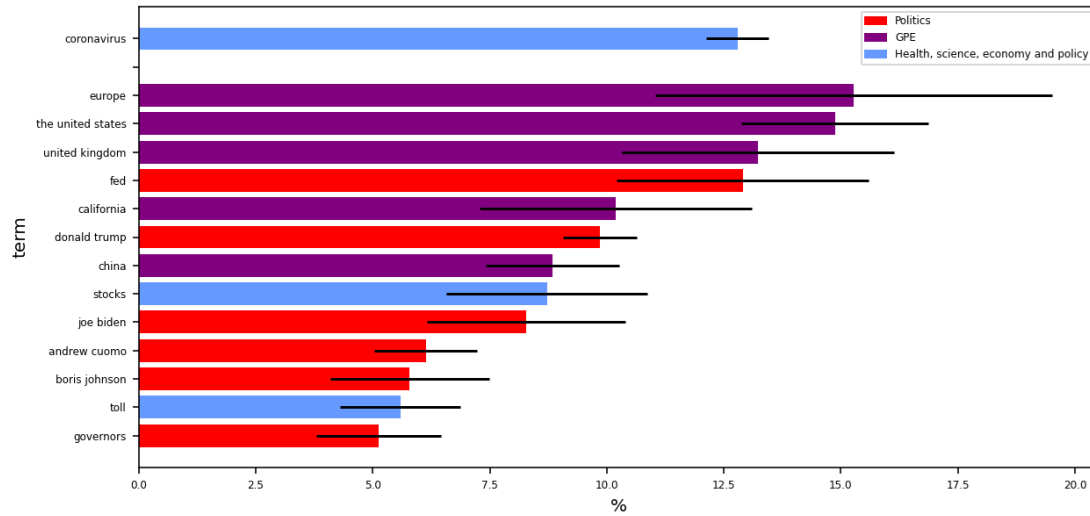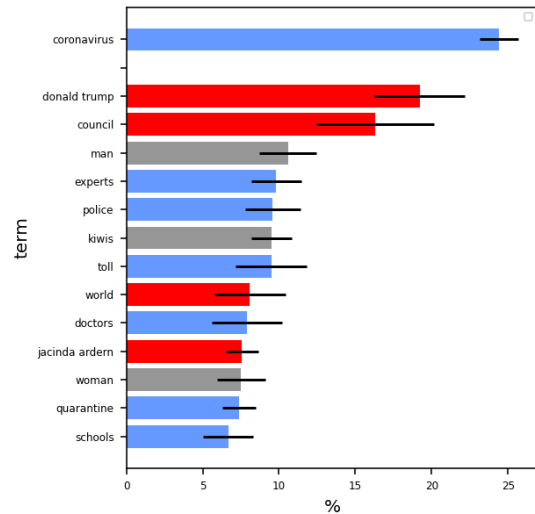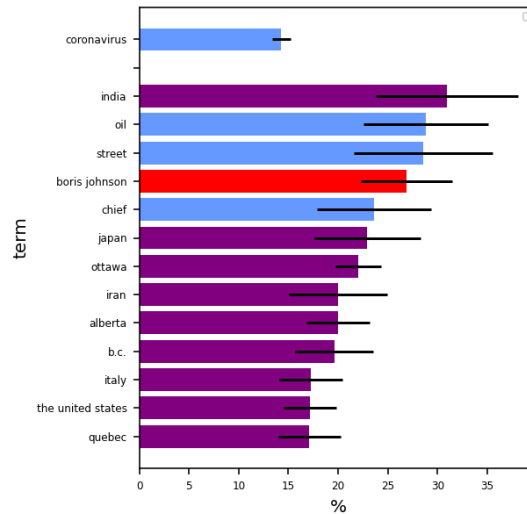
Fig. S1: **Share of Most Popular Grammatical Subjects in COVID-19 Articles**. Colors depict term categories: political terms in red; geopolitical entities in purple; policy issues related to health, science and economy in blue; grey for other.

Fig. S2: **Title to Body Ratio of Grammatical Subjects in COVID-19 Articles**. Colors depict term categories: political terms in red; geopolitical entities in purple; health, policy, science and economy in blue; grey for other. The size of the bubble is proportional to the share of articles where the term appeared as a subject in their title out of the articles where the term appeared as the subject of one or more of the sentences in their body.

Figure S3 further substantiates the high levels of politicization in COVID-19 coverage, and the focus on political personas, in the *NYT*. The figure compares the coverage of president Trump and PM Trudeau in the *NYT* and in the *G&M*. The US president was a source of media fascination around the world during the crisis (as well as in many other periods), and thus it is not surprising that he received extensive coverage in the G&M, indeed almost as extensively as the Canadian head of government himself. However, the gap between the US and Canadian outlets is clear: coverage in the *NYT* is dramatically more focused on the political leader of the USA, much more

than the Canadian coverage of the US leader and the Canadian leader alike This difference is statistically significant. The only exception—like in Figure 6 in the article—is in the grammatical subjects of image titles, where PM Trudeau is more frequently the grammatical subject than president Trump.
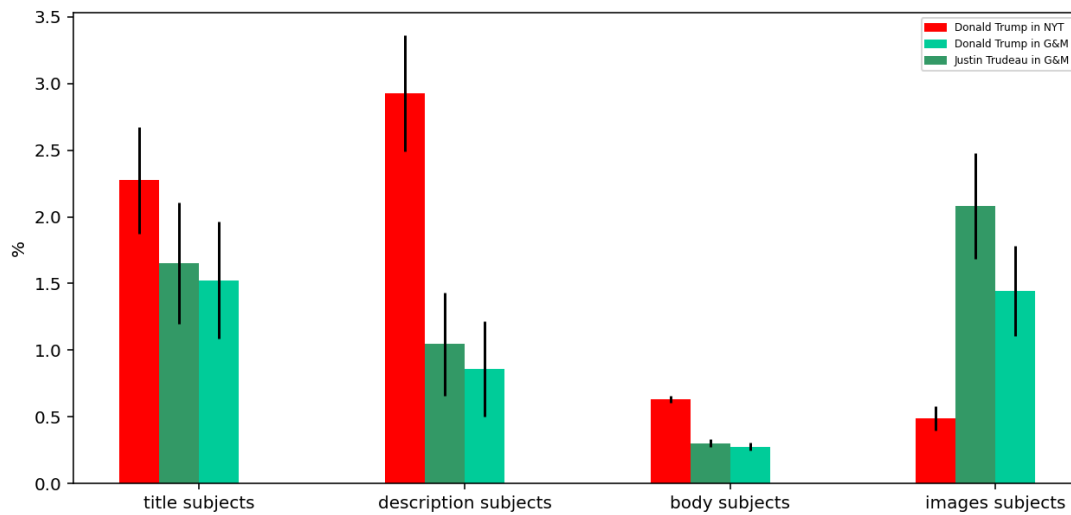

Fig. S3: **Comparative Share of Head of Government in Grammatical Subjects in *NYT* and *G&M*.** Comparing the share of leaders' appearance as grammatical subject in the title, description, body and image captions. Black lines indicate 2 std confidence intervals.