The Media and Foreign Powers: Does Market Access Matter for News Reporting?*

HENG CHEN
University of Hong Kong
LI HAN
ESSEC Business School

June 20, 2025

Abstract. Does news media coverage of autocracies hinge on their relationship with those regimes? Exploiting a large-scale media crackdown in May 2019 in China, in which multiple influential UK- and US-based news sites were blocked, we find that news outlets adopted a more negative tone in their coverage of China and reported more frequently on sensitive topics such as human rights, after being blocked, compared to those with no access change. Such effects are absent in news on economic topics and opinion articles. We investigate several mechanisms underlying these findings, including reduced self-censorship, diminished journalistic resources, and changes in readership composition after losing access.

Keywords: Media; Authoritarian Regime; Market Access; Word Embedding; Topic Modeling; Computational Linguistics

JEL Classification. L82 L88 F69 P00

*Heng CHEN: The Faculty of Business and Economics, The University of Hong Kong; Email: hengchen@hku.hk. Li HAN: Department of Economics, ESSEC Business School; Email: li.han@essec.edu.
This research project is partly funded by the General Research Fund of the Research Grants Council of Hong Kong (Project No. HKU 17500421 and Project No. HKUST 16508524) and by the Seed Fund for Basic Research of Hong Kong University (Project No. 201811159002).

1. Introduction

News media affect the beliefs and attitudes of voters and influence their decisions.[1] In the era of globalization, voters increasingly demand information about foreign countries, particularly regarding issues that intersect with domestic politics—such as import competition, technology rivalry, and climate change—and foreign policy concerns such as human rights and cross-border conflicts.

The economic integration of nondemocratic countries presents an attractive market to globally operating media outlets. For media organizations, market access in authoritarian countries holds strategic value beyond immediate profits. These outlets invest in cultivating their presence and influence in such markets, viewing them as potential future revenue bases, even when current returns are limited.[2] However, it also opens a door for economically important autocracies to influence media and, by extension, political decisions in democracies. Market access can become a source of leverage that authoritarian governments wield over foreign media.[3]

Does news media coverage of autocracies depend on their relationships with those regimes? While existing studies have documented how commercial interests, partisan preferences, and domestic government interference shape news content, we know little about how foreign autocratic governments may influence media coverage through market access. This gap in our understanding exists largely because the relationship between market access and media content is typically endogenous—news outlets' access to markets often depends on their coverage choices, making it difficult to isolate the causal effect of market access on reporting.

In this paper, we address this identification challenge by exploiting an unexpected shock to foreign media's market access in China—a large-scale "rectification" campaign launched in mid-2019, in which a batch of major international news outlets were blocked. This crackdown followed the breakdown of US-China trade negotiations, when China allegedly backtracked on a draft agreement, prompting President Trump to increase tariffs on Chinese products. While Chinese state media acknowledged the

[1] DellaVigna and Kaplan (2007) document the impacts of exposure to news reporting by Fox News. Enikolopov, Petrova, and Zhuravskaya (2011) show that access to independent news sources changed voting behaviors in Russia. La Ferrara, Chong, and Duryea (2012) show the impact of exposure to soap operas on fertility choices. Several other prominent studies on this issue include Strömberg (2004), Gentzkow and Shapiro (2004), Gentzkow (2006), Gerber, Karlan, and Bergan (2009), and Prat (2018).

[2] This strategic importance is exemplified by Facebook founder Mark Zuckerberg's efforts to gain market access in China through engagement with Chinese regulators. Media organizations recognize the value of cultivating audiences and strengthening their brands in foreign markets, particularly in populous and rapidly growing economies (see Smith 2017 and Paul and Sheera 2018).

[3] A case in point is Vietnam, a rapidly growing authoritarian country that has explicitly pressured Facebook and Google into censorship by threatening market exclusion (see Ratcliffe 2020 and Pearson 2020).

collapsed trade deal, the government sought to control information about its underlying causes, such as potential leadership disagreements and economic uncertainties, which could threaten social stability. Notably, the Chinese government blocked news websites based on their influence in China rather than their specific content, providing a unique setting to examine how market access affects media coverage .[4]

Our sample consists of two groups of English-language media outlets. The treatment group includes six major US and UK outlets that had significant presence in China before being blocked in late May 2019. The control group comprises 18 leading outlets—including top-circulation US newspapers, those listed in Baker, Bloom, and Davis (2016), and major UK national newspapers—whose access status in China remained unchanged (either consistently blocked or unblocked) during our study period. From these outlets, we collected articles published between January 2018 and April 2020 that mention at least one of our keywords: China, Chinese, Hong Kong, HongKonger (HongKongese), Russia, Russian, Iran, or Iranian.[5] The resulting corpus comprises over 1 million articles. We identify China-related articles using various criteria, from simple keyword counts to more sophisticated classification methods. While our main sample consists of articles containing at least three mentions of China-related keywords, we test the robustness of our findings by applying our empirical design to several alternative samples constructed using different selection criteria.

To systematically analyze reporting strategies across diverse content, we examine both extensive and intensive margins of coverage—frequency and tone, respectively. We focus particularly on news tone because it provides a comparable metric across time, outlets, topics, and articles, offering at least a conservative measure of media's adjustments in China-related coverage.

To measure article tone, we employ the Global Vectors for Word Representation (GloVe) algorithm to generate word embeddings and apply the "seed-word" approach of Rheault, Beelen, Cochrane, and Hirst (2016) to assess word sentiment. This method leverages the insight that words with similar sentiment cluster together in the embedded vector space–positive words cluster near other positive words, and likewise for negative words. We then aggregate these word-level sentiment scores to construct article-level tone measures for our primary analysis.

Having constructed measures for news reporting, we implement our empirical

[4]Gallagher and Miller (2021) document that the Chinese state's approach to online news control prioritizes the influence of information providers over content, as providers with substantial reach may challenge the state's dominance in public discourse. Our findings in Sections 2.1 and 3.1 align with their conclusions, demonstrating that blocked media outlets are characterized by greater influence and broader presence, while their news coverage of China is not more negative than that of unblocked media outlets.

[5]For all media outlets in our sample, we individually scraped articles containing our specified keywords from each outlet's website, where they maintain complete archives of their published content.

strategy to identify the causal effect of market access on media coverage. Specifically, we employ a difference-in-differences (DID) design that compares changes in reporting strategies on China between blocked and unaffected outlets before and after the campaign, controlling for outlet and year-month fixed effects.

Our difference-in-differences estimates reveal that the blockage significantly affected media outlets' reporting strategies. Specifically, news articles published by treated outlets exhibited more negative tone after the 2019 blockage, relative to control outlets. This effect is confined to news content, as we find no significant changes in the tone of opinion pieces.[6] The result remains robust across various alternative measures and samples constructed using different criteria.

We address a set of potential concerns about our findings. First, given our relatively small number of outlets, we validate our statistical inference using cluster-adjusted wild bootstrapping and randomization inference approaches. The negative blockage effect on news tone remains statistically significant under these alternative inference methods. Furthermore, our results are robust to iteratively excluding individual media outlets. This finding shows that our result is not driven by the chilling effect of the never-blocked media.

Second, we consider whether our findings might be driven by changes in reporting among never-blocked media responding to the crackdown. We show that our results hold when excluding never-blocked outlets from the sample, and that these outlets' tone did not diverge significantly from always-blocked media after the crackdown.

Third, we address concerns about endogeneity and pre-existing trends. Our results remain robust after excluding articles about potential triggers of the crackdown—the Sino-US trade negotiations breakdown, the anniversary of Tiananmen incident, and the 2019 Hong Kong protests. Moreover, an event study analysis reveals no differential pre-trends between treatment and control groups, with tone changes coinciding precisely with the crackdown timing, corroborating that the crackdown is not endogenous to the news content.

As a final set of checks, we address concerns about potential time-varying outlet-specific confounding factors. One might worry that the treated media became generally more responsive to authoritarian-related events after the crackdown. To mitigate this concern, we first demonstrate that our results remain robust after excluding news coverage of major events such as the COVID-19 crisis. We then implement a difference-in-differences-in-differences (DDD) design using Russia- and Iran-related articles as an additional comparison group. This analysis reveals no systematic changes in treated

[6] Most news media have an opinion section featuring articles with subjective views, including opinions, letters from readers, op-eds, and contributions from columnists. Because the editorial operation is independent from that of news sections, we examine news and opinion articles separately.

outlets' coverage of other authoritarian regimes after the crackdown, reinforcing our interpretation that the observed effects stem directly from the China market access shock.

To further unpack these effects, we employ a Latent Dirichlet Allocation (LDA) topic model to identify 12 distinct news topics. Our analysis reveals systematic heterogeneity in media response across topics: after being blocked, treated outlets adopted more negative tones and increased their coverage of politically sensitive topics, particularly human rights, compared to control outlets. However, for nonsensitive topics such as markets and economic growth, treated and control outlets showed similar trends in both tone and coverage frequency.

Our findings suggest that media coverage becomes more negative when outlets' relationships with Chinese authorities deteriorate. Several mechanisms may account for this effect, as we discuss in Section 7. First, prior to being blocked, outlets may have self-censored and moderated their tone on China-related issues to maintain market access. The crackdown effectively removed these self-imposed constraints, consistent with both anecdotal accounts and systematic evidence provided in this paper.

Second, blocked outlets may have reduced their journalistic resources for China coverage, leading to more opinion-based and less fact-based reporting. Using Chinese-language platforms as a proxy for pre-crackdown resource investment, we find that outlets that are less likely to adjust their journalistic resources exhibited smaller negative shifts in coverage tone following the blockage. These mechanisms–reduced self-censorship and journalistic resources–likely operated simultaneously, with their relative importance varying across outlets.

Finally, we explore alternative mechanisms, including changes in readership composition and potential retaliatory responses from blocked media. Using search-based attention proxies, we find that shifts in audience composition do not primarily explain the observed tone changes. The persistence of the negative tone effects suggests that factors beyond short-term retaliation drive our results.

For autocratic economic powers, our findings underscore the dilemma of accommodating foreign media. On the one hand, it is legitimate, from the point of view of the regime, to worry about foreign media's influence on citizens' information diet (Chen and Yang 2019; Cantoni, Chen, Yang, Yuchtman, and Zhang 2017).[7] On the other hand, autocratic regimes lose the strings that they can pull when foreign media are completely

[7] Chen and Yang (2019) designed an experiment in which Chinese students were incentivized to consume news from The New York Times and study such consumption's influence on the beliefs of the participants. In general, autocratic regimes understand that political information and narratives are important in shaping citizens' attitudes and therefore exert tight control over the information citizens are exposed to (Cantoni, Chen, Yang, Yuchtman, and Zhang 2017).

shut out.

Therefore, our study is related to a small body of literature on the influence of foreign media. Garcia-Arenas (2016) documented the impact of Radio Liberty on the 1991 Russian presidential elections and stressed the role of free media on regime change. Gagliarducci, Onorato, Sobbrio, and Tabellini (2020) study how BBC radio coordinated and mobilized Italian resistance forces during Nazi occupation.[8] We provide a new angle and study whether news content provided by free media may be affected by their commercial interests in autocratic countries.

Our paper adds to studies of the influence of governments on news media.[9] Existing research has focused on the role of domestic governments. For instance, Besley and Prat (2006) show that governments may use direct or indirect financial incentives to suppress news.[10] McMillan and Zoido (2004) provide evidence from Peru consistent with the direct channel. Di Tella and Franceschelli (2011) study the media market in Argentina and document that the government uses indirect channels such as government advertising to reduce negative coverage of government misconduct. Gentzkow, Petek, Shapiro, and Sinkinson (2015) show that party control of state governments did not influence the operations of partisan daily newspapers from 1869 to 1928, while Qian and Yanagizawa-Drott (2017) find that the Reagan and Bush Sr. administrations indeed influenced media outlets.[11] In particular, Simonov and Rao (2022) show that an authoritarian government can influence the ideological beliefs of citizens by investing in the quality of the government-controlled media platform and nonpolitical news content. Our paper shows that autocratic governments could also influence news businesses based in democracies.[12]

Furthermore, our study contributes to a growing literature in economics and political science that takes advantage of state-of-the-art techniques in computational

[8]In addition, DellaVigna, Enikolopov, Mironova, Petrova, and Zhuravskaya (2014) show that cross-border nationalistic Serbian radio provoked hatred toward Serbs in Croatia.

[9]This line of study is part of the literature in economics examining the determinants of news coverage. See an excellent survey by Prat and Strömberg (2013). Recent examples include analyses by Groseclose and Milyo (2005), Gentzkow and Shapiro ((2006) and (2010)) and Larcinese, Puglisi, and Snyder (2011).

[10]Economic leverage is also wielded by private enterprises to pressure news media to curtail unfavorable reporting about them. Germano and Meier (2013) theorize about this self-censorship mechanism of news media. On the empirical side, Beattie, Durante, Knight, and Sen (2021) show that auto manufacturer recalls are less extensively covered by newspapers in which the firms advertise more regularly.

[11]Other mechanisms have been studied in non-US contexts. For example, Stanig (2015) documents the impact of the defamation law wielded by Mexican governments in relation to news media. Durante and Knight (2012) provide evidence that the news content offered by the public television corporation in Italy shifted to the right when the elected government was center-right.

[12]Our study is also related to research on how access to news sources can distort news coverage. Ozerturk (2020) theorizes how access to politicians or governments may be used by these sources to extract more favorable press coverage, and Dyck and Zingales (2003) provide supporting evidence. The mechanism studied in our paper differs in that news outlets compromise their reporting to maintain access to a market for their products.

linguistics.[13] Our paper applies the word embedding approach to construct a measure of the negativity of news articles that cover a broad range of news events. Specifically, we utilize an algorithm proposed by Rheault, Beelen, Cochrane, and Hirst (2016) that measures the tone of parliamentary speeches in the UK. Gennaro and Ash (2022) use the embedding approach to quantify the use of emotion and reason in political discourse. Furthermore, Hansen, McMahon, and Prat (2018) and Catalinac (2016) both apply topic modeling, LDA in particular, to study political economy issues. Part of our analysis relies on topic modeling to uncover the underlying themes in the news corpus so that our definitions of various news topics are not excessively arbitrary.

2. Background and Research Questions

2.1. Media Environment in China and the 2019 Crackdown

China's relationship with openness and reform has followed a distinct trajectory: After reaching a peak with WTO entry in 2002, reforms aimed at increasing personal freedoms plateaued and then reversed course in the 2010s. This regression is particularly evident in media control and internet censorship. By 2020, China's internet censorship spending exceeded $6.6 billion USD. The government has become increasingly aggressive in policing online content, routinely deleting sensitive materials, and blocking entire news websites. *The New York Times*, for instance, has been blocked since 2012 after reporting on CCP leaders' family wealth.

According to the Foreign Correspondents' Club of China (2019), 23% of international news organizations' websites with China-based journalists are blocked, rising to 31% for English-language outlets. While VPNs offer a potential workaround, their use has been severely restricted since 2017 regulations banned unlicensed VPNs, with *Freedom House* noting increasingly sophisticated blocking techniques targeting VPN services.

One dramatic episode is China's "rectification" campaign to "clean up" its internet in May 2019. As reported by Reuters, this sweeping effort blocked or closed numerous websites and social media accounts, extending beyond political content. Notable casualties included *Wallstreetcn.com* (a prominent Chinese financial news site unrelated to the *Wall Street Journal*) and *Wikipedia*. The campaign also targeted Western media, blocking access to major news outlets from the United States, United Kingdom, Germany, Australia, and Singapore, including the *Washington Post* and *The Guardian*. The pattern of blockage suggests that the selection criterion was primarily based on outlets' prominence among Chinese readers—all blocked outlets were among the most frequently

[13] Among prominent examples of related studies, Gentzkow and Shapiro (2010) construct a media slant index based on partisan language used by the media. Shapiro, Sudhof, and Wilson (2020) develop a new sentiment-scoring model that accurately measures sentiment in economic news.

cited foreign news sources within China and had substantial Chinese readership prior to the ban—rather than being determined by their specific coverage patterns.

2.2. A Moving Red Line: US-China Trade Talks Upended

The crackdown was intended to control information on the unexpected breakdown of trade negotiations between the US and China.[14] The prolonged trade talks showed promising signs at the end of April 2019, when a draft trade agreement was crafted in high-level trade talks but took an abrupt turn on May 3, when the US negotiation team reported to "Washington [that] Beijing [had backtracked] on almost all aspects of the draft trade pact."[15] President Trump responded by escalating the trade war, increasing tariffs on US$200 billion worth of Chinese products from 10% to 25%, effective from May 10. While official Chinese media reported the collapse of the trade deal itself, discussion of its causes and consequences became sensitive. The government sought to suppress speculation about internal leadership disagreements, miscalculations of Trump's position, and growing economic uncertainty from deteriorating Sino-US relations. The trade war quietly became a new red line for media, largely unnoticed by foreign journalists in China. Some journalists suspected that the trigger to the crackdown is the impending 30th anniversary of the Tiananmen Incident (Waterson 2019).

Analysis of *Baidu Search Index* data–China's equivalent to *Google Trends*–reveals telling patterns about public attention during this period. As shown in Figure 1(a), search intensity for trade war-related topics spiked dramatically in May 2019, coinciding precisely with the media crackdown. In contrast, searches for other potentially sensitive topics (1989/Tiananmen Incident, Hong Kong, and Xinjiang) either peaked at different times or showed no significant variation.[16]

Our analysis of media coverage frequency aligns with public search patterns. Figure 1(b) shows the monthly mentions of key terms (trade war, 1989, Hong Kong, and Xinjiang) in our news sample (we elaborate on its construction in section 3.1). The parallel spikes in both media coverage and search activity suggest these metrics were

[14] The state media claimed this campaign targeted "illegal and criminal actions" and failure to protect personal information, but the timing–coinciding with the 30th anniversary of the Tiananmen Incident–led foreign journalists to suspect it was meant to restrict coverage of this event. However, later reports by Hong Kong media revealed broader motivations: controlling information about trade tensions with the U.S. For instance, wallstreetcn.com was allegedly shut down for translating Trump's May 2019 tweets about increasing tariffs after failed trade talks.

[15] For a summary of the key events of the trade negotiations, see Timmons (2020).

[16] On the Baidu search engine, the keyword "Tiananmen" is less informative than "1989" for the Tiananmen Incident, given that the location itself is also a site for military parades and tourism. The Baidu search index results for the keyword "Tiananmen" remained stable in the period until early October 2019, when they surged dramatically. This timing coincides with the military parade for the 70th anniversary of the People's Republic of China. Other phrases directly related to this incident are banned.

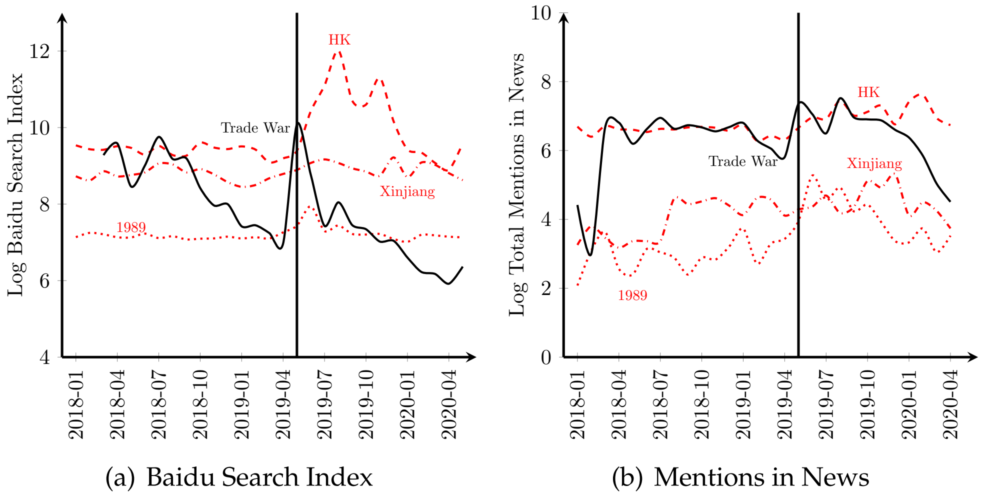


Figure 1. *Baidu Search Index and Monthly News Mentions by Issue. Panel (a) shows search patterns on Baidu: "trade war" searches peaked in early May 2019 following the US-China trade deal collapse; "1989" searches increased in early June 2019; "Hong Kong" searches surged in early August amid escalating protests; while "Xinjiang" searches remained stable throughout. Panel (b) displays monthly mentions of these issues in our news sample. The parallel patterns between search behavior and media coverage intensity suggest both responded to key events, particularly the trade deal breakdown.*

responding to the same underlying events. This correlation, particularly regarding trade war coverage, helps explain the unprecedented scope of the media crackdown.

2.3. Do Foreign Media Value Their Presence in China?

The Chinese government employs multiple tools to influence foreign media reporting, from business pressure to journalist obstruction.[17] Our study focuses specifically on market access, as it is both quantifiable and crucial to media outlets' strategic decisions. Major international outlets' significant investments in Chinese-language content — despite various restrictions — demonstrate the high value they place on maintaining market access. For example, the *New York Times*, *Wall Street Journal*, *Washington Post*, and *Reuters* as well as *Guardian* have gone out of their way to establish Chinese versions of their websites or translate their news to make them easily accessible to Chinese readers.

Media outlets value Chinese market access both for immediate commercial benefits and potential future opportunities, anticipating possible changes in the political climate — a perspective widely shared among news producers. For example, Craig Smith, a former *New York Times*'s Shanghai bureau chief and China managing director, once stated this calculation explicitly, reflecting on the situation prior to the outlet's 2012 blockage:

> "Our traffic ... grew nearly 70 percent last year alone. The New York Times

[17] See FCCC (2017) for details.

brand now has a firm foothold in the country and among the global Chinese diaspora. When news media restrictions relax, and I believe they eventually will, the Times's Chinese audience will most certainly take off."[18]

A media outlet's ban status affects its influence beyond direct readership through two key indirect channels: citations by Chinese media and social sharing. Both official media and individuals face penalties for citing or sharing content from banned foreign sources, substantially diminishing these outlets' potential impact.

3. Data

3.1. Sample Construction

Our analysis covers January 2018 to April 2020, spanning the June 2019 media crackdown. The sample includes articles from 24 major US and UK news outlets. Our treatment group consists of English-language news websites blocked during the crackdown, including the *Washington Post*, *NBC News*, *Huffington Post*, *Breitbart News*, *Guardian*, *Daily Mail*.[19]

Our control group comprises major English-language news outlets whose access status in China remained unchanged during our sample period. It includes three categories: (1) nine of the ten highest-circulation U.S. newspapers (*New York Times*, *Wall Street Journal*, *Boston Globe*, *Chicago Tribune*, *Los Angeles Times*, *News-Day*, *New York Post*, *Star Tribune*, and *USA Today*); (2) additional influential newspapers (*San Francisco Chronicle*, *Miami Herald*, and *Dallas Morning News*) identified by Baker, Bloom, and Davis (2016); and (3) UK-based outlets (*BBC News*, *Daily Mirror*, *The Telegraph*, *Financial Times*, *The Times*, and *The Independent*) to maintain geographic balance for UK treatment group.[20]

As shown in Table 1, our control group of 18 news outlets includes four "always-blocked outlets" (*New York Times*, *The Times*, *Wall Street Journal*, and *Financial Times*) blocked before 2018, and fourteen "never-blocked outlets" that maintained access throughout our study period.

As discussed in Section 2.1, the 2019 media crackdown has targeted outlets based on their influence. Using the Baidu search index as a proxy for influence in China, Figure

[18] See Smith (2017) for details.

[19] We exclude news sites blocked during this campaign that are based outside the US and the UK such as the *Straits Times of Singapore*. We verified the blocked status using information released by *GreatFire.org*, a nongovernmental organization that the FCCC partnered with to analyze and investigate foreign media access in China (discussed in section 2.1). Several independent testing services, such as *Chinese Firewall Test*, can verify the access status from China for any website.

[20] See Cision (2019) for the rank of media.

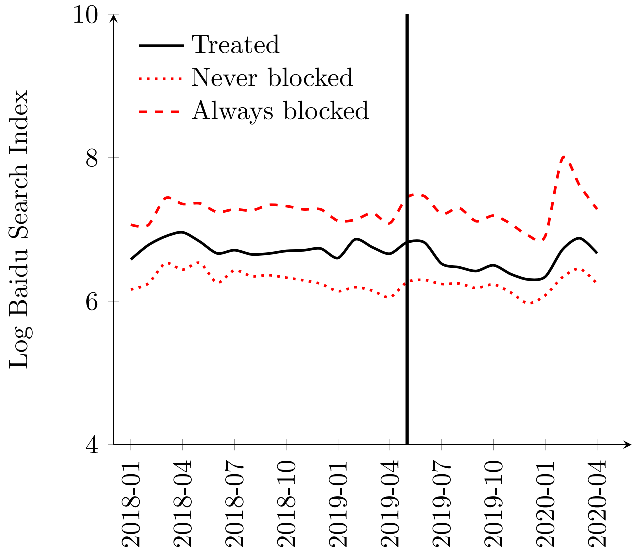


Figure 2. *Average Baidu Search Index by group. Chinese internet users search for the names of always-blocked media outlets most often, even though the outlets have been blocked. The media outlets newly blocked during the 2019 crackdown were searched for more often than the never-blocked outlets. The index for each group increased in February 2020, likely indicating that people searched for foreign media-reported information about the COVID-19 pandemic.*

2 shows that always-blocked outlets generated the highest search volume, followed by outlets blocked in 2019, with never-blocked outlets showing the lowest search frequencies.[21]

For our analysis, we collected articles containing China-related keywords ("China," "Chinese," "Hong Kong," "Hong Kongese," and "Hong Konger(s)"). Our main sample includes articles that mention these keywords at least three times—a transparent criterion that minimizes arbitrariness in sample selection. To assess robustness, we construct several alternative samples. First, we vary the keyword frequency threshold to create two additional samples: a "large sample" requiring only one keyword mention (to minimize exclusion of China-related articles) and a "small sample" requiring at least five mentions (to ensure stronger focus on China-related content).

Second, we create three refined samples by applying additional filters to our main sample: (i) a sample excluding articles that mention other countries in their headlines, (ii) an expanded sample that includes all articles from news outlets' dedicated China sections (see Table 2), and (iii) an expanded sample that adds articles mentioning China in their headlines. These alternative sample constructions help address potential concerns about article relevance and topical focus while demonstrating the robustness

[21] Like Google search trends elsewhere, Baidu search volume reflects public interest and engagement with specific topics or outlets. As China's primary search engine, Baidu data provides a reliable measure of an outlet's visibility and influence in the Chinese information space, with higher search frequencies indicating greater public attention and potential impact.

Table 1. *News Outlets*

Treatment	Control
Breitbart News	The New York Times* #3, blocked by 2012
Daily Mail	The Times, blocked before 2018
The Guardian*	The Wall Street Journal* #2, blocked by 2018
Huffington Post	Financial Times*, blocked before 2018
NBC News	Independent
The Washington Post* #6	The Boston Globe #10
	Chicago Tribune #9
	The Dallas Morning News
	Los Angeles Times #5
	Miami Herald
	Newsday #8
	New York Post #4
	San Francisco Chronicle
	Star Tribune #7
	USA Today #1
	BBC News*
	Telegraph
	Daily Mirror

Note: This table presents news outlets classification. Treatment outlets maintained China access post-2018 but lost it after the crackdown; control outlets were blocked before 2018 or never blocked. Major outlets (NYT, Times, WSJ, FT) were blocked in earlier years as shown. * indicates outlets with Chinese websites/translations. # shows U.S. circulation ranking (2018).

of our findings across different sampling criteria.

We classify articles into three mutually exclusive categories based on section designations: news (objective reporting and analysis), opinions (commentaries and columns), and entertainment (arts, lifestyle, and sports). We exclude entertainment articles from our analysis.

Our main analysis uses the "news sample"—news category articles from the China sample—spanning sections listed in Table 2 (Asia, Business, China, Education, Energy, Finance, Health, News, Politics, Technology, and World). This sample contains 47,711 articles (87.4% of total coverage), with 20,098 articles (36.8%) from the treatment group and 27,613 articles (50.6%) from the control group.

We also analyze the "opinion sample" of China-related opinion articles. Table 2 shows this comprises 6,883 articles (12.6% of total coverage), with 3,196 from the treatment group and 3,687 from the control group.[22, 23]

[22]Editorials by news staff are excluded as they likely reflect opinions rather than facts. However, the number of China-related editorials during our study period was minimal (fewer than 30), and their inclusion does not affect our results.

[23]The distribution of articles across categories varies between treatment and control groups partly

Table 2. Category and panel

	Control	Treatment	Total
Asia	3873 (7.1%)	1504 (2.8%)	5377 (9.8%)
Business	7782 (14.3%)	2563 (4.7%)	10345 (18.9%)
China (section)	487 (0.9%)	240 (0.4%)	727 (1.3%)
Education	49 (0.1%)	41 (0.1%)	90 (0.2%)
Energy	374 (0.7%)	206 (0.4%)	580 (1.1%)
Finance	1804 (3.3%)	134 (0.2%)	1938 (3.5%)
Health	346 (0.6%)	702 (1.3%)	1048 (1.9%)
News	3871 (7.1%)	8315 (15.2%)	12186 (22.3%)
Politics	2347 (4.3%)	3255 (6.0%)	5602 (10.3%)
Technology	1584 (2.9%)	565 (1.0%)	2149 (3.9%)
World	5096 (9.3%)	2573 (4.7%)	7669 (14.0%)
News Subtotal	27613 (50.6%)	20098 (36.8%)	47711 (87.4%)
Opinions	3687 (6.8%)	3196 (5.9%)	6883 (12.6%)
Total	31300 (57.3%)	23294 (42.7%)	54594 (100.0%)

Note: This table presents the distribution of articles across categories and treatment status in our sample. Categories are based on news outlets' own section classifications. News Subtotal aggregates all news categories (Asia through World), representing 87.4% of the sample, while Opinions (12.6%) captures editorial content. Treatment group includes outlets that maintained China access post-2018 but lost it after the crackdown; control group comprises outlets blocked before 2018 or never blocked. Numbers in parentheses show percentages of total sample (N=54,594).

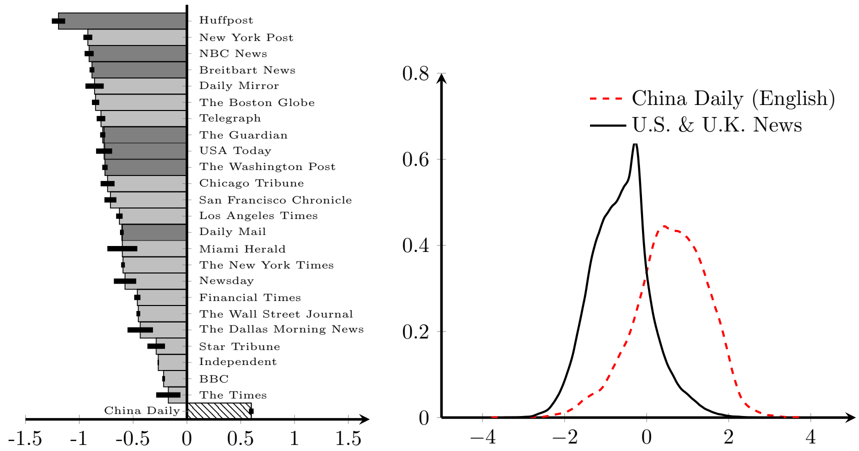
3.2. Measuring Negativity towards China

To measure article tone, we develop a corpus-based sentiment dictionary through a three-step process: 1) generating numerical vectors (embeddings) for each word in the corpus, 2) assigning emotion or tone scores to words using a sentiment lexicon, and 3) aggregating scores to the article level. This approach offers two key advantages. First, it is unsupervised and minimizes human input. Second, the vectorization process is context-adaptive: the vectors capture word meanings specific to how they are used in our news corpus.[24] This feature is particularly valuable for our study, as words may carry different emotional valences across contexts (e.g., parliamentary speeches, Wikipedia, news media) or across different time periods in news coverage.

While Appendix A provides full methodological details, we outline our key approach here. First, we create a vector space model that converts our corpus vocabulary into numerical vectors using the global vectors for word representation (*GloVe*) algorithm (Pennington, Socher, and Manning 2014), which encodes word meanings through word-word co-occurrence probability ratios.

due to differences in outlets' classification systems. For instance, similar content about China's environmental policies might be categorized under Asia, News, or Politics depending on the outlet. While this inconsistency limits panel-level comparisons across outlets, it does not affect our analysis of the overall news sample.

[24] This context-sensitivity helps avoid common dictionary-based approach limitations, such as difficulty handling polysemes and incomplete synonym coverage.



(a) Average tone by media outlet (b) Distribution of article tone

Figure 3. *Validity at the Press Level. The left panel displays the means and confidence intervals of article-level tone scores (the main measure) of UK- and US-based outlets in our news sample and those of China Daily (English news articles). The right panel shows the article-level score distribution for outlets in our news sample and that for China Daily.*

Second, we measure word tone using Rheault, Beelen, Cochrane, and Hirst (2016)'s algorithm. The approach calculates how similar each word is to a set of clearly positive words (like "excellent" or "wonderful") versus a set of clearly negative words (like "terrible" or "awful"). These reference words, called seed words, are carefully chosen to have unambiguous meanings and avoid domain-specific terms. A word receives a higher tone score if it is more similar to positive seed words and less similar to negative ones. The similarity between words is measured by how often they appear in similar contexts in our corpus.

Third, we develop three article-level tone measures. Our main measure is the average score of all words in an article (excluding stop words). To address potential noise from mixed sentiments about China versus context, we create a China-based score using only words from sentences containing "China" or "Chinese." For robustness, we construct a nonneutral score using only words with strong emotional valence (excluding words within one standard deviation of the lexicon's mean score).[25]

To validate our tone measures, we employ multiple strategies. Our first validation compares English-language articles from *China Daily* — the Chinese government's official English-language newspaper and primary international mouthpiece — against our sample of US and UK media outlets. This comparison provides a clear reference point, as *China Daily*'s state-directed coverage should differ systematically from inde-

[25]One might be concerned that the frequency of mentions of "China" or "Chinese" mechanically affects the sentiment score. This is not the case in our analysis. Given that the average article length is approximately 800 words, and neither our sentiment score nor the main results would be changed much when we exclude those mentions from calculations.

pendent Western reporting. For this analysis, we include all articles that mention China at least three times, ensuring comprehensive coverage of China-related content across both sources. Figure 3 confirms this expectation: the left panel reveals China Daily's consistently positive average article tone contrasts sharply with the negative averages across all Western outlets, while the right panel shows China Daily's article-level tone distribution is distinctly skewed toward positive values compared to our sample outlets. For additional validation, Appendix A demonstrates strong correlation between our automated tone measures and human-coded ratings, using *New York Times* articles as representative examples.

3.3. Summary Statistics

Table A1 presents summary statistics for our key variables. In the news sample, the default tone scores average -0.73 and -0.54 for treatment and control groups respectively (columns 1-2), with a difference of 0.19 (column 3, standard errors clustered at press level). The China-based scores are more negative, averaging -0.82 for treatment and -0.59 for control groups. When excluding neutral words within one standard deviation of the mean (nonneutral score), tone scores become substantially more negative, averaging -1.59 for treatment and -1.34 for control groups. For comparison, columns 4-6 show similar patterns in the opinion sample, with consistently more negative scores in the treatment group across all measures. Article length, measured by logged word count, is similar between treatment and control groups in the news sample (6.24 vs. 6.19), but shows a larger difference in the opinion sample (6.65 vs. 6.13).

4. Identification Strategy

As discussed in section 2.1, the large-scale crackdown in May 2019 was based on the influence of news outlets rather than the content published by specific outlets, and intended to control information on and attention to the unexpected breakdown of trade negotiations. This consideration motivates our use of a difference-in-differences (DID) model to identify how losing access to China affected the media's handling of China-related articles. We start by comparing changes in the tone toward China of the treated outlets with those of the control outlets using the following specification:

$$y_{ipjt} = \beta^{DID} \left(Treated_j \times Post \right) + X_i \gamma + \rho_p + \mu_j + \lambda_t + \epsilon_{ipjt} \quad (1)$$

where y_{ipjt} is the measure of the tone of article i in panel p published by outlet j at time (in month) t. T_j is the indicator for the treatment group; *Post* is a dummy variable that takes the value of 1 if article i was published in or after June 2019 and is 0 otherwise, and X_i is a vector of article-level control variables, including the total word count and the total number of occurrences of words "China" and "Chinese" and occurrences of words "Taiwan" and "Taiwanese" in article i, all in logarithm forms. These variables capture article i's length and relevance to China, respectively. We include panel, outlet, and month fixed effects — denoted by ρ_p, μ_j and λ_t, respectively — to control for panel-, outlet- and time-specific factors that affect the tone of news articles. The inclusion of these fixed effects renders the dummy variables T_j and $Post$ redundant in this regression. All standard errors are clustered at the press level.

The key coefficient of interest is β^{DID} in Equation (1), which captures the impact of the 2019 blockage on outcome variables. We attribute a significant estimate of β^{DID} to losing market access under the parallel trends assumption that the treated media outlets would have followed a trend of the outcome variables parallel to that of the control outlets had they not been blocked in 2019.

The first challenge to our research design is that the number of outlets in our sample is relatively small, especially that of treatment outlets. The within-outlet correlations may lead to an underestimation of standard errors. To address this concern, we report three sets of p values adjusted for this bias. First, we follow the suggestion by Bertrand, Duflo, and Mullainathan (2004) to report the cluster-correlated Huber-White standard errors for all specifications. Second, we report p values computed using the cluster-adjusted wild bootstrap (WB) method, following MacKinnon and Webb (2018) and considering each press as a cluster. Third, we also report p values based on the randomization inference (RI) test (Rosenbaum 2002).[26] Both WB and RI approaches yield conservative estimates. If the respective p values are sufficiently small, the over-rejection problem caused by the small number of clusters should not be a serious concern.

Another potential threat to our identification strategy is that the blocking decision might have been endogenous to outlets' news content or pre-existing content trends. To address this concern, we conduct robustness tests by excluding articles mentioning likely triggers of the crackdown—the US-China trade war, the Tiananmen Incident, and the Hong Kong protests—thereby isolating the treatment effect on broader China coverage unrelated to these sensitive topics.

Next, we test whether the treatment outlets had developed an increasingly harsher tone over time before the blockage compared to the control group using an event study

[26]We construct the sampling distribution of the estimated $\hat{\beta}$ by repeatedly randomly assigning the treatment outlet and estimating the placebo effects. The p value is computed by noting where our estimated effect lies in the distribution of placebo effects.

model specified as follows:

$$y_{ipjt} = \sum_{\tau=-17,\tau\neq-1}^{10} \alpha_\tau (Treated_j \times Month_\tau) + X_i\gamma + \rho_p + \mu_j + \lambda_t + \mu_{ipjt}. \quad (2)$$

Using May 2019 as the base period, we examine sixteen months before and eleven months after the crackdown. We define $Month_\tau$ ($\tau = -17, ..., 11$) as monthly dummies from January 2018 to April 2020, with $\tau = 0$ representing June 2019. The coefficients α_τ of interaction terms $T_j \times Month_\tau$ should not differ significantly from zero for $\tau < 0$ if pre-trends are parallel. If the blockage induced harsher coverage, we expect negative α_τ for $\tau \geq 0$.

To address potential "chilling effects"–where unblocked media might self-censor after the crackdown—we conduct two analyses. First, we reestimate our models using only always-blocked outlets as controls. Second, we perform a placebo test comparing never-blocked outlets (as treatment) to always-blocked outlets (as controls) using a similar DID specification to Equation (1). If unblocked media were chilled into reducing negativity, this placebo test should yield positive effects. The absence of such effects would validate our control group construction.

Another identification concern is that outlets' specializations might drive differential responses to post-blockage events involving authoritarian politics. We address this through two approaches. First, we verify robustness by excluding articles about major post-blockage events (e.g., COVID-19 crisis). Second, we compare treated outlets' coverage of China with their coverage of Russia and Iran—chosen as comparable authoritarian states receiving significant media attention. Using a combination of our China sample and similarly constructed Russia and Iran samples (see Table A8, Appendix B.10 for the summary statistics), we estimate the following specification:

$$y_{ipct} = \beta^C \left(China_i \times Post\right) + X_i\gamma + \rho_p + \nu_c + \mu_j + \lambda_t + \epsilon_{ipct}, \quad (3)$$

where y_{ipcjt} denotes tone of article i published by press j in panel p on country c in month t. $China_i$ is a binary indicator that equals 1 if the article covers China, and ν_c represents country fixed effects. A negative β^C would suggest the tone changes stem from China's crackdown rather than outlets' general coverage of authoritarian politics.

To address confounding from time-varying group-specific factors (e.g., general tone shifts toward authoritarian regimes), we further implement a triple-difference (DDD) model combining China, Russia, and Iran coverage across treated and control outlets:

$$\begin{aligned} y_{ipcjt} &= \delta_1 \left(Treated_j \times Post\right) + \delta_2 \left(Treated_j \times China_i\right) + \delta_3 \left(China_i \times Post\right) \\ &+ \beta^{DDD} \left(Treated_j \times China_i \times Post\right) + X_i\gamma + \rho_p + \nu_c + \mu_j + \lambda_t + \epsilon_{ipcjt}, \quad (4) \end{aligned}$$

where y_{ipcjt} measures tone for article i in panel p about country c by outlet j at month t. The coefficient β^{DDD} captures how China coverage tone differences between treated and control outlets changed post-blockage, relative to the changes in the Russia/Iran coverage tone differences between treated and control outlets. A statistically insignificant β^{DDD} would suggest our baseline DID estimates reflect general shifts in authoritarian regime coverage rather than China-specific effects.

5. Does the Market Access Matter for News Reporting?

5.1. Baseline Results

How did the news outlets change their tone after losing access to the Chinese market? Column (1) of Table 3 presents the estimation results for the DID model with main effects and without any controls. Column (2) adds controls and panel fixed effects. The statistically insignificant coefficient on the treatment group indicator (*Treated*) suggests comparable pre-crackdown tone between treatment and control media. The statistically significant negative coefficient on *Post* indicates a general tone hardening across all outlets.

Column (3) includes the Month- and Press-fixed effects. The key interaction term (*Treated* $\times$ Post) shows that treatment outlets' average tone score decreased by 0.155 relative to control outlets after the blockage, significant at the 1% level. This effect's magnitude is substantial: the gap between our news sample's average tone (-0.73) and China Daily's (0.44) is 1.1, and the blockage effect represents approximately 15% of this difference. In other words, losing market access caused treatment outlets' tone to diverge from China Daily's by an additional 15% compared to control outlets.

To examine whether our estimate is subject to the over-rejecting problem caused by the small number of clusters, we show the p values of the effect computed using cluster-adjusted wild bootstrap (WB) and randomization inference (RI) in the square and curly braces, respectively, for each specification in Table 3. For the news sample, the WB-based p values are 7.9%, 6.2% and 2.1% for the three respective specifications, while the RI-based p values are 6.8%, 3.9% and 7.1%. Both sets of p values corroborate the robustness of our DID estimates.

To further eliminate the possibility that a particular outlet drives our findings, we reestimate Equation (1) by excluding one media outlet at a time. The result, reported in Table A3 of Appendix B.3, remains robust.

How do these effects differ between news and opinion content? Columns (4)-(6) of Table 3 reveal a striking contrast in the opinion sample. Unlike news coverage, opinion pieces show significant negative main effects for both treatment group and

Table 3. *Baseline Difference-in-Differences Results: Tone Changes*

	Outcome Variable: Article-level Tone					
	News Sample			Opinions Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated $\times$ Post	-0.120*	-0.164**	-0.155***	0.086	0.087	0.021
	(0.069)	(0.062)	(0.053)	(0.074)	(0.073)	(0.059)
[*WB p-value*]	[0.079]	[0.062]	[0.0218]	[0.620]	[0.580]	[0.878]
{*RI p-value*}	{0.068}	{0.039}	{0.071}	{0.795}	{0.798}	{0.693}
Treated	-0.093	-0.004		-0.232***	-0.257***	
	(0.108)	(0.094)		(0.064)	(0.060)	
Post	-0.274***	-0.249***		-0.215***	-0.216***	
	(0.042)	(0.037)		(0.037)	(0.036)	
Controls	No	Yes	Yes	No	Yes	Yes
Month FE	No	No	Yes	No	No	Yes
Press FE	No	No	Yes	No	No	Yes
Panel FE	No	Yes	Yes	No	No	No
R-Squared	0.069	0.128	0.197	0.045	0.048	0.128
N	47,711	47,711	47,711	6,883	6,883	6,882

Notes: This table shows baseline estimates of the crackdown effect on news tone. The dependent variable is the article-level tone score. Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. Controls include log of total word count and log of China-related terms ('China', 'Chinese', 'Taiwan'). Sample includes articles with ≥ 3 China-related keywords. Standard errors (in parentheses) clustered at media outlet level; * p<0.1, ** p<0.05, *** p<0.01. *P*-values from wild bootstrap [square brackets] and randomization inference {curly braces}.

post-blockage period, indicating consistently harsher criticism of China from treatment outlets and an overall negative shift across all outlets after the crackdown. However, the interaction term (*Treated* $\times$ *Post*) remains statistically insignificant across all specifications, even with media and month fixed effects in column (6). The wild bootstrap and randomization inference-based p-values for the opinion sample estimates exceed 50%, confirming the absence of significant blockage effects in this sample. This suggests that while opinion pieces were generally more critical, the blockage did not cause differential changes in tone between treatment and control outlets — a notable departure from the news results. This pattern aligns with media outlets treating opinion pieces as individual authors' perspectives rather than institutional positions.[27]

To further validate our findings, we conduct a placebo test using entertainment articles. The results in Table A2 indicate no significant effects of the blockage on these articles, regardless of control variables and fixed effects. This strengthens our confidence that the reporting changes were specific to news articles.

[27] For instance, newspapers routinely publish diverse viewpoints on contentious issues to inform readers about existing views. It has long been a practice and a tenet in journalism that there is a "wall" between the news and opinion sides of business; i.e., reporters working for the news section and those working for opinion sections remain independent (see Kovach 2021).

5.2. Robustness Tests

Crackdown endogenous to news content? To examine whether the crackdown was endogenous to news content, we investigate whether articles mentioning the trade war, the Tiananmen Incident, or Hong Kong—the suspected triggers of this crackdown—drive our results. We re-estimate Equation (1) by excluding articles containing each of these terms separately. The results, reported in columns (1)-(3) of Table A4 in Appendix B, show consistently significant negative effects (-0.169, -0.155, and -0.166 respectively, all significant at the 1% level). Even with substantial sample reductions (for instance, the exclusion of articles containing "Hong Kong"' reduces the sample to 36,277 observations), the identified blockage effects on news tone remain robust and similar in magnitude to our baseline estimates.

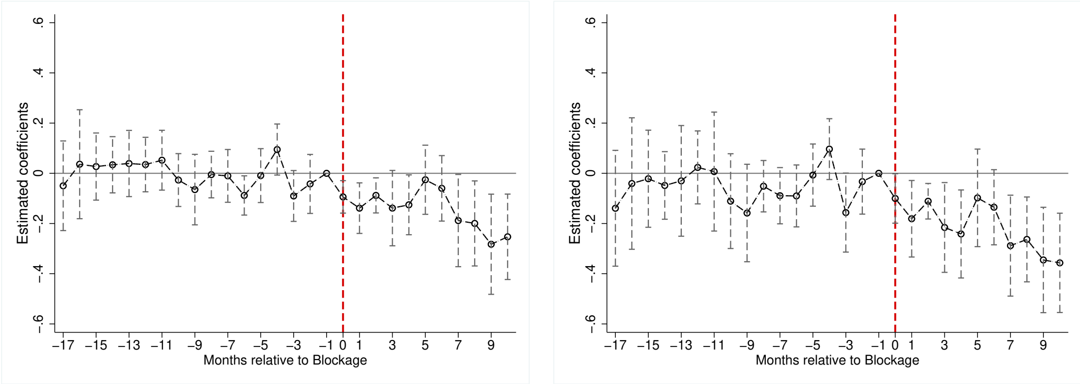
Preexisting trends in news content? We use the event study model to examine the time at which the trends in tones in the treatment and control groups diverged. We estimate Equation (2) using our benchmark tone scores as the outcome variable. Figure 4(a) illustrates the estimated coefficients α_τ (versus the number of months relative to the blockage) and their 95% confidence intervals.

Except for α_{-6} that is significant, the estimated coefficients α_τ are overall statistically insignificant for $\tau < 0$, indicating no difference in pre-trends between the treatment and control groups before the blockage. This finding rules out the concern that the treated outlets were blocked in May 2019 because they exhibited an increasingly negative tone toward China.

In contrast, starting from June 2019 (the month immediately after the blockage), the estimated coefficients α_τ are consistently negative and significant except for α_5 and α_6. In other words, articles from the treated media outlets exhibited a greater deterioration in tone than those in the control group. The timing of this divergence coincides precisely with the crackdown waged by the Chinese government, suggesting that the effect arises from the response of treated outlets to the blockage.

Furthermore, we conduct the honest test for parallel trends based on the smoothness restrictions following Rambachan and Roth (2023). Figure A3(a) in Appendix B shows that we can reject a null effect unless the post-treatment violations will deviate from a linear extrapolation of the pre-trend violations by 0.06. The result shows that our DID estimates are fairly robust to the violation of the parallel trends.

Chilling Effects? Does our result arise because the never-blocked outlets in the control group responded to the crackdown by adopting a more positive tone towards China? To explore this, we reestimate the same event study model of Equation (2) using only always-blocked outlets as the control group. The pattern, illustrated in Figure 4(b), is



(a) Control Group I: All (b) Control Group II: Always-blocked

Figure 4. *Event Study Model. The left panel (a) illustrates coefficients and the associated confidence intervals estimated with the event study model and by using all outlets in the control group. The right panel (b) illustrates the respective coefficients resulting from using the always-blocked outlets as the control group, i.e., control group II. The patterns in both estimations are rather similar. There is no difference in the preexisting trends between the treatment and control groups before the blockage. The timing of the divergence between the treatment and control groups coincides precisely with the crackdown. The month between before the crackdown is treated as the base period.* $Month_\tau$ *(where* $\tau = -17, \ldots, 10$*) represents dummy variables for the months from January 2018 to April 2020. In particular,* $\tau = -1$ *indicates the month of May 2019, at the end of which the crackdown occurred.*

rather similar to that for the entire control group shown in Figure 4(a), indicating that it is not driven by a potential chilling effect. We conduct the sensitivity test based on the smoothness restrictions following Rambachan and Roth (2023). Figure A3(b) in Appendix B shows that we can reject a null effect unless we are willing to allow for the linear extrapolation across consecutive periods to be off by more than 0.08, further demonstrating the robustness of our results.

We further test whether the never-blocked media outlets responded to the crackdown differently from always-blocked outlets, which did not respond. We perform a placebo test by relabeling the always-blocked media as the control group, and the never-blocked media as the pseudo-treatment group. Using the sample for only these two groups of media outlets, we estimate Equation (1) for a variety of measures of news tone and observe no significant blockage impact on the never-blocked media. The result, shown in Table A5 in Appendix B, reassures us that there was no significant chilling effect and that our construction of the control group is valid.

Different responsiveness to the COVID-19 pandemic? Could the harsher tone have resulted from treated outlets being inherently more responsive to major newsworthy events that occurred after the blockage? Of particular concern is the COVID-19 pandemic, which began in late January 2020 and persisted throughout 2021. Two pieces of evidence suggest that this is not the case. First, our event study shows that the divergence in tone between treatment and control groups emerged in the months before the pandemic, indicating that COVID-19 coverage was not the primary driver of this effect. Second, we re-estimate Equation (2) by excluding articles containing COVID-19-related keywords. The temporal pattern of coefficients (shown in Figure A4 in Appendix B) closely mirrors the results from our full sample.

Different responsiveness to authoritarian politics? Could the treatment effect reflect the inherently different reporting of treated media on authoritarian regimes or foreign affairs? We address this by comparing China-related coverage to Russia- and Iran-related coverage within treated media outlets using Equation (3). Column (1) of Table 4 shows two key findings. First, treated media initially adopted a more positive tone toward China than toward Russia/Iran (coefficient on *China*: 0.592, significant at the 1% level). Second, these outlets specifically increased negativity toward China after the crackdown (*China* $\times$ *Post*: -0.266, significant at the 1% level), while their coverage of Russia and Iran did not show similar changes. Column (2) shows that this pattern remains strong with the inclusion of the press and month fixed effects, suggesting that our main results are not driven by the shift in treated outlets' general approach to covering authoritarian regimes.

One may worry that the increased hostility toward China reflects a general trend in media attitudes toward authoritarian countries, which could potentially confound our DID estimate. To explore this, we estimate the DDD model (4) by combining China, Russia and Iran samples, with results reported in columns (3) and (4) of Table 4. The significant negative coefficients of the triple interaction *Treated* $\times$ *China* $\times$ *Post* suggest that the blockage led the treated media increased their negativity toward China relative to Russia and Iran, compared to the control media. It is worth noting that the DDD estimates of the blockage effect are approximately -0.15, a magnitude very close to the DID estimate.

Interestingly the insignificant coefficients on the interaction *Treated* $\times$ *China* show that the treated media had no particular bias against China before the blockage, while the insignificant coefficients on the interaction *Treated* $\times$ *Post* reveal that the treated media's coverage of Russia and Iran did not diverge from that of the control media (columns (3) and (4)). In summary, while all media became increasingly negative toward these authoritarian regimes, the treated media exhibited additional negativity specifically toward China after the blockage.

Robustness to alternative measures and samples. We assess the robustness of our results using alternative measures and samples. First, we estimate Equation (1) using alternative sentiment measures: the China-based scores and the nonneutral scores (columns (1) and (2) of Table A6). Both estimates remain significantly negative and consistent with our baseline result from Table 3. We then re-estimate the DID model

Table 4. Russia and Iran samples as a comparison group

	Outcome Variable: Article-level Tone			
	Treatment Media with China, Russia and Iran Samples		All Media with China, Russia and Iran Samples	
	Difference-in-Differences		Triple Differences	
	(1)	(2)	(3)	(4)
China×Post	-0.266***	-0.287***	-0.132***	-0.144***
	(0.040)	(0.039)	(0.026)	(0.031)
China	0.592***		0.496***	
	(0.058)		(0.064)	
Post	-0.141***		-0.122***	
	(0.012)		(0.039)	
Treated			-0.068	
			(0.108)	
Treated ×China×Post			-0.149**	-0.155***
			(0.060)	(0.053)
Treated × Post			-0.030	-0.017
			(0.040)	(0.024)
Treated ×China			0.065	0.096
			(0.098)	(0.105)
Controls	Yes	Yes	Yes	Yes
Press FE	No	Yes	No	Yes
Month FE	No	Yes	No	Yes
Panel FE	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	Yes
R-Squared	0.185	0.228	0.218	0.306
N	37638	37638	86295	86295

Notes: This table presents estimates using Russia and Iran coverage as additional comparison groups. Controls include log of: total word count; occurrences of 'China'/'Chinese', 'Taiwan', 'Russia'/'Russian', and 'Iran'/'Iranian' terms. Standard errors (in parentheses) clustered at media outlet level; * p<0.1, ** p<0.05, *** p<0.01. Columns 1-2 show difference-in-differences estimates using only treatment media, while columns 3-4 present triple differences estimates using all media. The sample includes articles containing relevant country-specific keywords for China, Russia, and Iran.

using two alternative samples based on keyword frequency: articles containing at least 5 or 1 China-related keywords (columns (3) and (4)). Both specifications yield estimates that closely align with our baseline finding, confirming that our results are robust to sample construction choices.

We re-estimate our model using three alternative sample construction methods discussed in Section 3.1: (i) a restricted sample that excludes articles mentioning other countries in their headlines, (ii) an expanded sample incorporating all articles from news outlets' dedicated China sections (see Table 2), and (iii) an expanded sample that includes articles with China mentioned in their headlines. Table A7 shows that all specifications yield significantly negative coefficients, consistent with our benchmark results. This consistency across different sample selection criteria demonstrates the robustness of our findings.

(a) Topic 1: Market (b) Topic 2: Trade (c) Topic 6: Rights

Figure 5. *Word Clouds for Key Topics in China Coverage. Word size indicates relative frequency of occurrence within each topic. Panel (a) shows frequently occurring terms related to market developments and economic growth; Panel (b) displays terms associated with trade; Panel (c) presents commonly used words in human rights discussions.*

6. News Topics: Intensive v.s. Extensive Margins

Having established changes in tone following market access loss, we now examine topic-specific adjustments. Using topic modeling, we analyze how outlets altered their reporting across different topics, examining the effects arising from both the intensive and extensive margins.

6.1. Intensive Margin: News Tone across Topics

To characterize topics or themes, we estimate an LDA topic model (Blei, Ng, and Jordan 2003) with our China news sample. LDA is a generative probabilistic model in which the assignment of words to topics and the assignment of topics to documents are jointly estimated. In this model, a topic is defined as a distribution over words; that is, the word probabilities for a given topic sum to one. A document is a distribution over topics; that is, the topic proportions across all topics for a document sum to one.[28] The details of our estimation are relegated to Appendix C.

The number of topics K is the key choice to make; it varies based on the study's purpose. For example, choosing a large number of topics, we would obtain topics such as China's relations with Japan, Europe and the UK. Choosing a smaller number of topics, we would obtain coarser topics such as China's foreign relations.

In our benchmark model, we select $K = 12$ topics, a number that allows for distinct representation of key themes relevant to our analysis while avoiding topic redundancy.

[28] LDA trades off two goals: (i) for each document, the algorithm allocates words to as few topics as possible, and (ii) for each topic, the algorithm assigns a high probability to as few words as possible. Therefore, topics (weighted word lists) emerge endogenously from the estimation without requiring pre-specified words to characterize them. Another output is a multinomial distribution over topics for each document (weighted topic lists).

The identified topics are clearly interpretable, spanning economic issues (market and growth, trade, Chinese companies), international relations (U.S. relations, North Korea, Taiwan, and Russia relations), social and political matters (human rights, social issues, Hong Kong protests), and COVID-related coverage (general reporting, travel restrictions, outbreaks). Tables A12 and A13 in Appendix C present the top keywords for each topic. These topics represent consistently covered themes in the media, with most topics (except those related to COVID-19) appearing both before and after the crackdown. For instance, the Hong Kong protest topic encompasses both the 2019 protests and various other political movements in Hong Kong from earlier years. Figure 5 displays word clouds for three representative topics: market and growth, trade, and human rights. The complete set of word clouds for all topics can be found in Figure A5 of Appendix C.

Based on the estimated likelihood of an article containing a specific topic, we create 12 subsamples, each of which consists of articles that are most likely to represent one particular topic. Specifically, for each topic $k = \{1, 2, \ldots, K\}$, we rank articles by each article i's probability of representing topic k, i.e., p_{ik}, and select articles from the top quartile. Since LDA allows each document (an article, in our case) to contain multiple topics, the subsamples are not mutually exclusive.

We estimate Equation (1) using each of the 12 subsamples. Figure 6 displays the estimated coefficients and their 95% confidence intervals for all topics.[29] For topics 1-4 (market and growth, trade, Chinese companies, and US affairs), the treated media outlets showed no differential response to the blockage compared to the control group. However, for the human rights topic (topic 6, presented in Table A15) — a consistently sensitive issue for the Chinese government — the blockage significantly intensified the negative tone in media coverage. We observe similar patterns of increased negativity in coverage of relations with North Korea, Taiwan and Russia (topic 7), social issues (topic 8), Hong Kong protests (topic 9), miscellaneous China-related news (topic 10).

Given that the COVID-19 crisis emerged after the crackdown, the DID estimates for COVID-related topics (topics 5, 11, and 12) are not interpretable. The model can still be estimated, as LDA might assign high COVID-topic probabilities to pre-COVID articles covering similar themes.[30] As shown in Section 5.2, our main findings remain robust after removing the COVID-19 coverage.

The first four topics cover Chinese economic issues and U.S. relations — areas traditionally considered within acceptable bounds of Chinese censorship. While trade

[29] Results from 1-5 topics are collected in Table A14, while the remaining topics are presented in Table A15. Both tables are relegated to Appendix C.

[30] For example, news articles about epidemic outbreaks in 2019 or earlier, which are unrelated to COVID-19, are given high probabilities of covering COVID-19-related topics. See Wee (2019) for details.

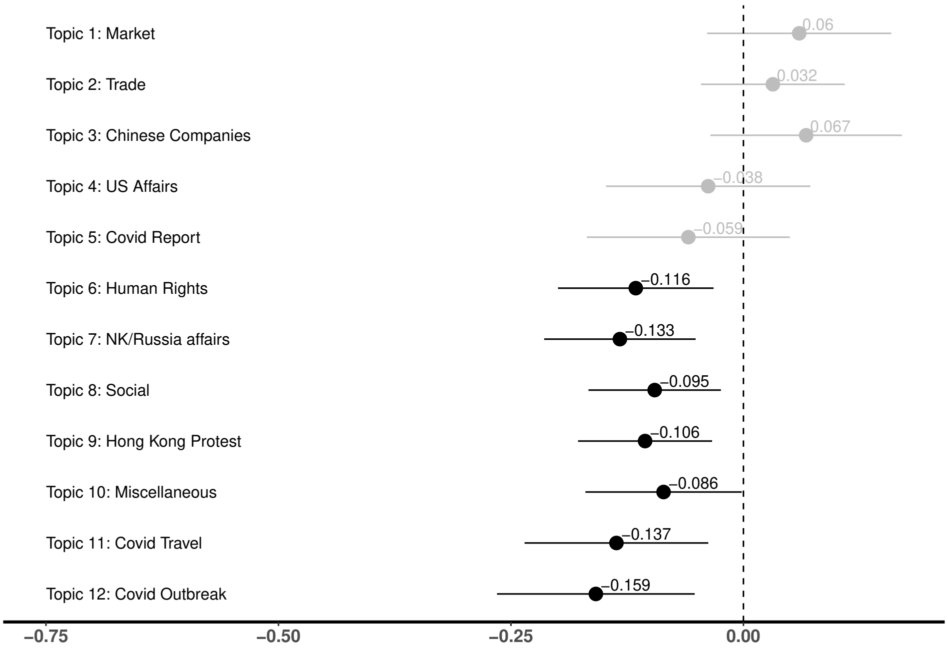


Figure 6. *Impacts of Media Crackdown at the Intensive Margin across Topics. We construct 12 subsamples by selecting articles in the top quartile of topic probability based on LDA topic modeling. The figure shows the DID estimates using Equation (1) and 95% confidence intervals (bold if the p-value is below 0.1) for twelve topics: Market, Trade, Chinese Companies, U.S. Affairs, COVID Report, Human Rights, NK/Russia affairs, Social, Hong Kong Protest, Miscellaneous, COVID Travel, and COVID Outbreak. Covid-related topics (5, 11, 12) are estimable but not interpretable due to pandemic timing. Controls include the log of total word count and the log of China-related terms ('China', 'Chinese', 'Taiwan'). Standard errors are clustered at the media outlet level.*

(topic 2) was initially non-sensitive, it became contentious following the breakdown in trade negotiations that preceded the crackdowns, as discussed in Section 2.2. Topics 6 through 10 encompass more politically sensitive subjects than economic ones, including human rights, military activities related to neighboring countries, and domestic social incidents (such as mine collapses and industrial accidents). Our findings indicate that the crackdown's impact on media tone was largely driven by these politically sensitive areas that typically draw scrutiny from Chinese authorities. To demonstrate that our main findings are not driven by any particular topic, we re-estimate Equation (1) by excluding articles of one topic at a time. The results are shown in Table A16. The results remain highly consistent with our main estimates across all topic exclusions.

6.2. Extensive Margin: Reporting Frequency across Topics

We next examine how the blockage affected media outlets' topic coverage frequency. To analyze this extensive margin, we classify articles into our 12 identified topics. We create a dummy variable A_{ik} that equals 1 if article i's probability of representing topic k (p_{ik}) falls in the top quartile across all articles (consistent with Section 6.1), and 0 otherwise. Using these classifications, we aggregate the number of articles per topic for

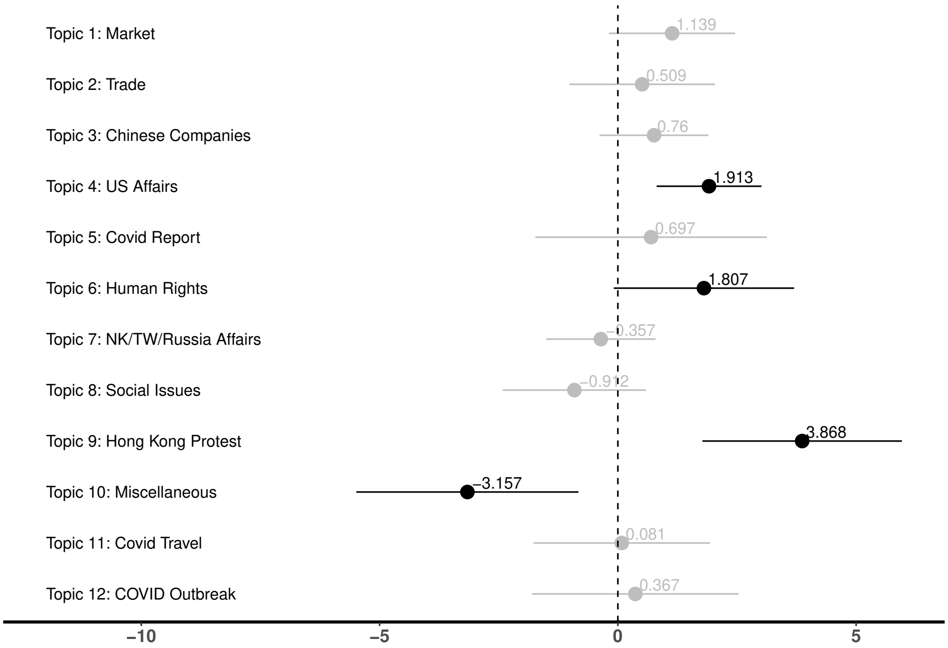


Figure 7. *Impacts of Media Crackdown at the Extensive Margin across Topics. Using the weekly article count per outlet in each topic as dependent variables, the figure shows difference-in-differences estimates and 95% confidence intervals (bold if the p-value is below 0.1) for twelve topics: Market, Trade, Chinese Companies, US Affairs, Covid Report, Human Rights, NK/Russia affairs, Social, Hong Kong Protest, Miscellaneous, COVID Travel, and COVID Outbreak. COVID-related topics (5, 11, 12) are estimable but not interpretable due to pandemic timing. Standard errors are clustered at the media outlet level.*

each media outlet weekly. Summary statistics for this analysis are presented in Table A17 in Appendix C.

We analyze changes in weekly article frequency per topic using a specification similar to Equation (1), estimated at the week-outlet level with week and outlet fixed effects. Figure 7 displays the estimated DID coefficients and their 95% confidence intervals across all topics.[31]

Our results show differentiated responses across topics. Coverage of Chinese economic topics (topics 1–3) showed no significant differences between treated and control groups post-blockage. However, treated outlets increased their weekly publication frequency relative to control outlets by approximately 2 articles for U.S. affairs and human rights topics, and 4 articles for Hong Kong political conflicts — all statistically significant changes. These findings suggest that while treated media intensified their coverage of sensitive topics, they did not increase coverage of non-sensitive topics. Notably, we find no significant change in the overall frequency of China-related reporting, which is consistent with our findings that they reduced coverage of miscellaneous topics by about 3 articles weekly.

[31] Detailed results are presented in Tables A18 (topics 1–5) and A19 (topics 6–13) in Appendix C.

7. Interpretations

Our findings suggest that the relationship between autocracies and the media in democracies is a relevant determinant of news coverage of those autocratic regimes. Following China's market access restrictions, outlets adopted more negative tones and increased their coverage of sensitive topics.

What could be behind the shift in their news reporting strategy? It is possible that multiple mechanisms are at play simultaneously, resulting in the observed changes. In this section, we explore several potential mechanisms that could drive these strategic adjustments in reporting behavior. Importantly, these mechanisms are not mutually exclusive and may operate in parallel — for instance, media outlets could simultaneously experience reduced pressure for self-censorship while facing new resource constraints after being blocked. The relative importance of different mechanisms is likely to vary across outlets based on their pre-blockage market position in China, institutional characteristics, and editorial strategies. With this framework in mind, we examine the evidence for each potential mechanism.

7.1. Censorship?

One interpretation of our findings is that prior to the loss of access, news outlets optimized and managed their reporting strategies by trading off their influence and profit at home and abroad in both the short and the long run. Fearing retaliation by Chinese censors in the case of crossing red lines, those outlets may have intentionally compromised their reporting strategy, such as softening how they report on China. Once access was lost, news outlets would have fewer constraints on choosing how and what to report. Anecdotes supporting this mechanism abound.[32]

Consistent with this interpretation, no change is found in the tone of opinion articles. Opinion articles are produced independently of news articles. News outlets typically include and publish contributions that prevent diverse or even contrasting views on the same issues. They are conventionally considered to reflect the authors' own views, for which the outlets claim no responsibility. It is reasonable that news outlets have much less or no incentive to interfere with their publication (Kovach 2021).

Our results on the differential effects across topics lend more support to the self-censorship interpretation. The Chinese government is known to be less tolerant of

[32] For example, according to Folkenflik (2020), Bloomberg's editor-in-chief justified this editorial decision in a private (but taped and eventually leaked) conference call with the outlet's China-based investigative team: "It is for sure going to, you know, invite the Communist Party to, you know, completely shut us down and kick us out of the country. So I just don't see that as a story that is justified." The editor went on in the same conference call to suggest a compromise strategy to address the dilemma at hand: "There's a way to use the information you have in such a way that enables us to report, but not kill ourselves in the process and wipe out everything we've tried to build there."

critical coverage of political issues (such as human rights) than of economic issues. The findings in Sections 6.1 and 6.2 suggest that the media intentionally toned down their negativity toward China and reduced the quantity of news content on sensitive topics before the blockage. These findings reveal that the media treat sensitive issues with extra caution when they have access to the Chinese market.

A complementary mechanism related to censorship suggests that direct experience with censorship heightens media outlets' sensitivity to issues of political freedom in China. This firsthand encounter with information control may sharpen journalists' and editors' focus on related topics in their coverage.

7.2. Journalistic Resources?

While self-censorship offers a coherent explanation for our findings, alternative mechanisms warrant consideration. Blocked outlets may have reduced their journalistic resources in China, potentially affecting coverage quality. Limited investigative capacity could lead to a greater reliance on opinion-based writing rather than fact-based reporting, as journalists struggle to access primary sources. Such shifts toward opinion-heavy coverage, which tends to be more emotional than factual reporting, could contribute to the observed changes in tone.

Journalistic resources are the combination of human capital (reporters, editors, and support staff on the ground), physical infrastructure (local offices and equipment), and established networks of local sources and contacts that enable comprehensive news gathering and fact-checking in China. Although we cannot directly measure journalistic resources, we proxy them through outlets' investments in Chinese-language platforms, including dedicated websites and regular content translation. If the cut in journalistic resources drove the tone change, outlets with greater pre-crackdown investments should show more resilience to the blockage's effects, as their *resource adjustments* were likely less drastic than those of others for several reasons.[33] Under this reasoning, we conduct a heterogeneity analysis of the blockade effect for media outlets with a Chinese platform and those without, by including the interactions between an indicator for having a Chinese platform ($Chinese\ platform$) and the variables $treated$, $Post$, and $treated \times Post$ in Equation (1). The result, presented Table A9 in Appendix C, supports this hypothesis: the positive and significant coefficient on the triple interaction term ($treated \times Post \times Chinese\ Platform$) indicates that outlets with stronger Chinese-

[33]These outlets could likely maintain more stable resource allocation for the following reasons: (1) they have made significant fixed investments in physical infrastructure and human capital, which would be costly to dismantle, (2) their Chinese-language operations often serve broader Asian markets beyond mainland China, making complete withdrawal less optimal, and (3) they typically have more diversified networks of sources and contacts, which can be maintained even under restricted access.

language presence experienced smaller shifts in coverage tone following the blockage.[34]

To further examine the relationship between journalistic resources and coverage tone, we develop an independent test using COVID-19 as an exogenous negative shock to journalistic resources.[35] The pandemic disrupted news gathering operations through travel restrictions, reduced access to sources, and limited on-the-ground reporting capabilities, particularly for foreign correspondents in China. Our hypothesis is that outlets with a Chinese-language platform, being larger and more invested in reporting on China, would experience fewer disruptions to their reporting and investigative capabilities due to the pandemic, resulting in less adjustment in news tone.

Using the sample after the media crackdown (i.e., from June 2019 to April 2020, spanning the COVID crisis), we compare the tone of coverage on China between outlets with and without Chinese-language platforms, before and after COVID-19, employing a specification similar to Equation (1). We replace the interaction term with Post COVID-19 $\times$ Chinese Platform, where Post COVID-19 indicates articles published after the COVID-19 crisis, and Chinese Platform indicates outlets with Chinese-language platforms. Table A10 shows that COVID-19 led to a significant decline in coverage (-0.269), but outlets with Chinese-language platforms exhibited a relatively positive differential effect (0.103 in the baseline specification and 0.137 with fixed effects, both statistically significant at the 5% level). This pattern supports our hypothesis that better-resourced outlets maintained more stable coverage during the crisis. By focusing on the post-blockage period, this test provides independent evidence of resource effects on news coverage, illuminating a mechanism driving our main findings on the impact of the crackdown.

7.3. Readership Composition and Grievances

Another potential mechanism is readership composition: blocked outlets might have adjusted their coverage to match American and British readers' preferences for negative China coverage after losing their Chinese audience. To test this hypothesis, we examine whether controlling for Chinese and non-Chinese readership explains the differential response between treated and control outlets.

While direct readership measures are unavailable, we construct attention proxies using search data. For Chinese reader attention, we use the monthly average of Baidu search index of each outlet's name (see Section 3.1, page 9). For Western attention to China coverage, we use the monthly average Google Trends data from U.K. and U.S.

[34]We corroborate this mechanism by examining foreign media influence in China, measured by Baidu search intensity in the Chinese domain. Outlets with above-median Baidu search index are considered to have higher influence, which typically correlates with greater journalistic resources in China. When we replace the Chinese platform indicator with this alternative measure of influence, we find similar results.

[35]We thank one of the referees for suggesting this test.

domains, specifically searching "newspaper name + China" (e.g., *The Washington Post China*). This refined search term helps isolate China-specific attention from general outlet readership.

Table A11 in Appendix B.8 presents DID estimates (Equation 1) with additional readership proxy controls and their interactions with the post-crackdown dummy. The *Treated* $\times$ *Post* coefficients remain similar to baseline estimates in Table 3, suggesting audience composition changes do not primarily drive the observed tone shifts.

Having found limited evidence for readership composition effects, we next consider emotional responses to the blockage: blocked outlets might have adopted harsher tones toward China in retaliation for market access loss. However, this interpretation faces empirical challenges. While grievances might explain short-term reactions, they would likely dissipate absent commercial benefits. Our event study model (Section 5.2) shows persistent effects that do not diminish over time, even when excluding COVID-19 coverage (Figure A4, Appendix B), suggesting factors beyond short-term retaliation drive the observed changes.

8. Concluding Remarks

It is not unlikely that free media that enjoy protection from the rule of law at home succumb to influence from authoritarian regimes abroad. This phenomenon is new, partly because it is only in recent decades that rising economic powers have been undemocratic yet so economically intertwined with democratic countries.

Autocratic governments' manipulation of or interventions in news production have recently become an important issue in political discourse. However, discussions have centered mainly on the impact of direct interventions, e.g., foreign governments may wage disinformation campaigns or seek to control news outlets that target audiences in democratic countries. We discover a less apparent channel through which news production could be influenced by foreign governments leveraging economic power. This channel may pose no less of a threat to the backbone of democracy than outright interventions, given its concealed nature.

The mechanism underlying our findings is not unique to the news industry. The Economist has recently observed that the global film industry is not free from meddling by Chinese censors. Since China is becoming the world's largest cinema market by revenue, even overtaking the U.S., Hollywood has geared its products to the Chinese market and, when necessary, altered films to please Chinese censors, including changing the versions for global audiences (The Economist 2020a and The Economist 2020b). The case of *Netflix* represents the other side of the coin, as it has never been allowed to enter the Chinese market, and therefore has had a free hand to commission documentaries about pro-democracy movements in Hong Kong, over which censors are fret.

Our findings also beget new thinking on the censorship strategy of autocrats. Dealing with foreign entities—be it *The New York Times* or *Hollywood*—is tricky. Allowing them to exert influence at home creates uneasiness for autocratic regimes. However, autocrats who have economic power at their disposal lose the strings that they can pull behind the scenes when foreign entities are shut out entirely. The optimal degree of openness may require trading off influence at home and abroad, which is an interesting topic for future research.

References

Baker, Scott R., Nicholas Bloom, and Steven J. Davis. 2016. "Measuring Economic Policy Uncertainty." *The Quarterly Journal of Economics* 131(4): 1593–1636.

Beattie, Graham, Ruben Durante, Brian Knight, and Ananya Sen. 2021. "Advertising Spending and Media Bias: Evidence from News Coverage of Car Safety Recalls." *Management Science* 67(2): 698–719.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics* 119(1): 249–275.

Besley, Timothy, and Andrea Prat. 2006. "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability." *American Economic Review* 96(3): 720–736.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *The Journal of Machine Learning Research* 3: 993–1022.

Cantoni, D., Y. Chen, D. Y. Yang, N. Yuchtman, and Y. J. Zhang (2017). Curriculum and Ideology. *Journal of Political Economy 125*(2), 338–392.

Catalinac, Amy. 2016. "From Pork to Policy: The Rise of Programmatic Campaigning in Japanese Elections." *The Journal of Politics* 78(1): 1–18.

Chen, Yuyu, and David Y. Yang. 2019. "The Impact of Media Censorship: 1984 or Brave New World?" *American Economic Review* 109(6): 2294–2332.

Cision. 2019. "2019 State of the Media Report" *Cision Research*.

DellaVigna, S., R. Enikolopov, V. Mironova, M. Petrova, and E. Zhuravskaya (2014). Cross-border media and nationalism: Evidence from Serbian radio in Croatia. *American Economic Journal: Applied Economics 6*(3), 103–32.

DellaVigna, Stefano, and Johannes Hermle. 2017. "Does Conflict of Interest Lead to Biased Coverage? Evidence from Movie Reviews. *Review of Economic Studies* 84(4): 1510–1550.

DellaVigna, Stefano, and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics* 122(3): 1187–1234.

Di Tella, Rafael, and Ignacio Franceschelli. 2011. "Government Advertising and Media Coverage of Corruption Scandals. *American Economic Journal: Applied Economics* 3(4): 119–51.

Durante, Ruben, and Brian Knight. 2012. “Partisan Control, Media Bias, and Viewer Responses: Evidence from Berlusconi’s Italy." *Journal of the European Economic Association* 10(3): 451–481.

Dyck, Alexander, and Luigi Zingales. 2003. “The Media and Asset Prices." Working Paper, Harvard Business School.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2011. “Media and Political Persuasion: Evidence from Russia" *American Economic Review* 101(7): 3253–85.

FCCC. 2017. “Access Denied: Surveillance, Harassment and Intimidation as Reporting Conditions in China Deteriorate" *Foreign Correspondents’ Club of China*.

Friedrich, R., M. Luzzatto, and E. Ash (2020). Entropy in legal language. In *NLLP 2020 Natural Legal Language Processing Workshop 2020. Proceedings of the Natural Legal Language Processing Workshop 2020 co-located with the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2020)*, Volume 2645, pp. 25–30. CEUR-WS.

Folkenflik, David. 2020. “Bloomberg News Killed Investigation, Fired Reporter, Then Sought To Silence His Wife." *NPR Report*.

Gagliarducci, S., M. G. Onorato, F. Sobbrio, and G. Tabellini (2020). War of the Waves: Radio and Resistance During World War II. *American Economic Journal: Applied Economics 12*(4), 1–38.

Gallagher, Mary E., and Blake Miller. 2021. “Who Not What: The Logic of China’s Information Control Strategy." *The China Quarterly* 248(1): 1020–1045.

Garcia-Arenas, J. (2016). The Impact of Free Media on Regime Change: Evidence from Russia. *The Quarterly Journal of Economics 128*, 105–164.

Gennaro, Gloria and Ash, Elliott. 2022. “Emotion and Reason in Political Language." *The Economic Journal* 132(643): 1037–1059. Oxford University Press.

Gentzkow, Matthew. 2006. “Television and Voter Turnout." *The Quarterly Journal of Economics* 121(3): 931–972.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. “Text as Data." *Journal of Economic Literature* 57(3): 535–74.

Gentzkow, Matthew, Nathan Petek, Jesse M. Shapiro, and Michael Sinkinson. 2015. “Do Newspapers Serve the State? Incumbent Party Influence on the US Press, 1869–1928." *Journal of the European Economic Association* 13(1): 29–61.

Gentzkow, Matthew, and Jesse M. Shapiro. 2006. "Media Bias and Reputation." *Journal of Political Economy* 114(2): 280–316.

Gentzkow, Matthew, and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence from US Daily Newspapers." *Econometrica* 78(1): 35–71.

Gentzkow, Matthew A. and Jesse M. Shapiro. 2004. "Media, Education and Anti-Americanism in the Muslim World." *Journal of Economic Perspectives* 18(3): 117–133.

Gerber, Alan S., Dean Karlan, and Daniel Bergan. 2009. "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions." *American Economic Journal: Applied Economics* 1(2): 35–52.

Germano, Fabrizio, and Martin Meier. 2013. "Concentration and Self-censorship in Commercial Media." *Journal of Public Economics* 97: 117–130.

Groseclose, Tim, and Jeffrey Milyo. 2005. "A Measure of Media Bias." *The Quarterly Journal of Economics* 120(4): 1191–1237.

Hansen, Stephen, Michael McMahon, and Andrea Prat. 2018. "Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach." *The Quarterly Journal of Economics* 133(2): 801–870.

Huang, Junming, Gavin Cook, and Yu Xie. 2021. "Do Mass Media Shape Public Opinion Toward China? Quantitative Evidence on New York Times With Deep Learning." *arXiv preprint arXiv:2012.07575.*

Kovach, Bill, and Rosenstiel, Tom. 2021. "The Elements of Journalism: What News-people Should Know and the Public Should Expect." 4th ed., Crown, New York, NY.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics* 4(4): 1–31.

Larcinese, Valentino, Riccardo Puglisi, and James M. Snyder . 2011. "Partisan Bias in Economic News: Evidence on the Agenda-setting Behavior of US Newspapers." *Journal of Public Economics* 95(9–10): 1178–1189.

MacKinnon, J. G. and M. D. Webb (2018). "The wild bootstrap for few (treated) clusters". *The Econometrics Journal 21*(2), 114–135.

McMillan, John, and Pablo Zoido. 2004. "How to Subvert Democracy: Montesinos in Peru." *Journal of Economic Perspectives* 18(4): 69–92.

Mozur, Paul and Frenkel, Sheera. 2018. "Facebook Gains Status in China, at Least for a Moment." *The New York Times*.

Ozerturk, Saltuk. 2020. "Media Access, Bias and Public Opinion." Working paper, Southern Methodist University.

Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. 2014. "Glove: Global Vectors for Word Representation." In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543.

Prat, Andrea. 2018. "Media Power." *Journal of Political Economy* 126(4): 1747–1783.

Prat, Andrea, and David Strömberg. 2013. "The Political Economy of Mass Media." *Advances in Economics and Econometrics* 2: 135.

Qian, Nancy, and David Yanagizawa-Drott. 2017. "Government Distortion in Independently Owned Media: Evidence from US News Coverage of Human Rights." *Journal of the European Economic Association* 15(2): 463–499.

Pearson, James. 2020. "Exclusive: Vietnam Threatens to Shut Down Facebook over Censorship Requests - Source." *The Reuters*.

Rambachan, Ashesh and Roth, Jonathan. 2023. "A More Credible Approach to Parallel Trends." *Review of Economic Studies* 90(5): 2555–2591.

Ratcliffe, Rebecca. 2020. "Facebook and YouTube Accused of Complicity in Vietnam Repression." *The Guardian*.

Rheault, Ludovic, Kaspar Beelen, Christopher Cochrane, and Graeme Hirst. 2016. "Measuring Emotion in Parliamentary Debates with Automated Textual Analysis." *PloS One* 11(12): e0168843.

Rosenbaum, Paul R. 2002. "Overt Bias in Observational Studies." In *Observational studies*, pp. 71–104. Springer, New York, NY.

Shapiro, Adam Hale, Moritz Sudhof, and Daniel J. Wilson. 2020. "Measuring News Sentiment." *Journal of Econometrics*.

Simonov, Andrey, and Justin Rao. 2022. "Demand for Online News under Government Control: Evidence from Russia." *Journal of Political Economy* 130(2).

Stanig, Piero. 2015. "Regulation of Speech and Media Coverage of Corruption: An Empirical Analysis of the Mexican Press. *American Journal of Political Science* 59(1): 175–193.

Strömberg, David. 2004. "Radio's Impact on Public Spending." *The Quarterly Journal of Economics* 119(1): 189–221.

Smith, Craig. 2017. "The New York Times vs. the 'Great Firewall' of China." *The New York Times*.

The Economist. 2020a. "How Hollywood Should Deal with Chinese Censors" *The Economist*.

The Economist. 2020b. "Hollywood's Chinese Conundrums" *The Economist*.

Timmons, Heather. 2020. "Timeline: Key Dates in the U.S.-China Trade War." *The Reuters*.

Waterson, Jim. 2019. "Chinese Government Blocks Guardian Website." *The Guardian*.

Wee, Sui-Lee. 2019. "Pneumonic Plague Is Diagnosed in China." *The New York Times*.

Online Appendix

(Not intended for publication)

A. Tone Construction

The *GloVe* Algorithm

In this study, the tone of each article is an aggregation of each word in the text. To determine the tone of each word, we need to represent its meaning. One of the techniques of meaning representation is word embedding, i.e., representing a word by a dense and low-dimensional numerical vector in a meaningful manner. Given that some form of meaning is encoded in those vectors, semantic relations between words can be captured by the geometry of corresponding vectors. This work uses the algorithm of Global Vectors for Word Representation (*GloVe*), proposed by Pennington, Socher, and Manning (2014), to perform word embedding, which is one of the leading algorithms that excel in word analogy accuracy. *GloVe* is at least as efficient as the SKIM and CWOB methods. The algorithm is widely used and has been cited by more than 19,000 scientific articles so far.

First, it is essential for the *GloVe* algorithm to build the word-word co-occurrence matrix X, inside which each entry X_{ij} represents the number of times word j occurs in the context of word i, where context is defined as a window centered around the focus word. Therefore, the probability that word j appears in the context of word i is constructed by:

$$P_{ij} = \frac{X_{ij}}{X_i},$$

where X_i is the number of times any word appears in the context of word i.

Second, two features distinguish the *GloVe* method from others. (i) It utilizes the "co-occurrence probabilities ratios" rather than the raw probabilities. Pennington, Socher, and Manning (2014) show that the co-occurrence ratios gather more information and better capture the relationship between words. (ii) An efficient and workable function F is proposed to predict those ratios– such that

$$F\left(w_i, w_j, \tilde{w}_k\right) = \frac{P_{ik}}{P_{jk}}, \tag{5}$$

where w_i and w_j are two word vectors and $\tilde{w}_k$ is a context word vector.

One leading and frequently cited example that the authors use to illustrate this insights is as follows: "ice co-occurs more frequently with solid than it does with gas, whereas steam co-occurs more frequently with gas than it does with solid. Both words co-occur with their shared property water frequently, and both co-occur with the unrelated word fashion infrequently. Only in the ratio of probabilities does noise from non-discriminative words like water and fashion cancel out, so that large values (much greater than 1) correlate well with properties specific to ice, and small values (much less than 1) correlate well with properties specific of steam. In this way, the ratio of probabilities encodes some crude form of meaning associated with the abstract concept of thermodynamic phase (The *GloVe* official site)."

Third, equation (5) associates word vectors on the left-hand side with text statistics (i.e., those co-occurrence probabilities ratios) on the right hand side. That is, while those word vectors are to be learned, the probability ratios are observable empirically. A cost/objective function is defined to capture the differences between them. The *GloVe* algorithm minimizes this objective function by learning meaningful word vectors representations.

The News Corpus, Training and Tone Construction

To generate meaningful word embeddings using *GloVe*, we constructed a large corpus by scraping over 1 million articles from all sample outlets (control and treatment) that mention China, Hong Kong, Russia, or Iran-related keywords at least once. This corpus contains nearly 800 million tokens. Using the authors' C source code, we trained the model with standard parameters: 15-word context windows and 300-dimensional word vectors. The output assigns a vector representation to each word in our corpus.

To measure the positivity/negativity of each word, we follow the algorithm proposed by Rheault, Beelen, Cochrane, and Hirst (2016). Their key insight is that words carrying positive sentiment tend to cluster closer to known positive words and further from negative ones in the vector space.

For implementation, we use their carefully curated seed words: 100 positive and 100 negative terms, specifically selected to avoid polysemants and analogies (listed in their appendix Tables H and I). We extract vector representations for these seed words from our trained corpus to anchor the sentiment analysis.

Next, the distances between words are constructed with cosine similarity of word vectors. The similarity between w_i and w_j is:

$$\frac{w_i w_j}{||w_i||||w_j||}$$

where $||w_i||$ is the norm of word vector w_i and the similarity is in a $[-1, 1]$ interval. Intuitively, completely irrelevant words give a similarity score close to 0; two closely located vectors w_i and w_j in the space lead to a similarity score close to 1; antonym

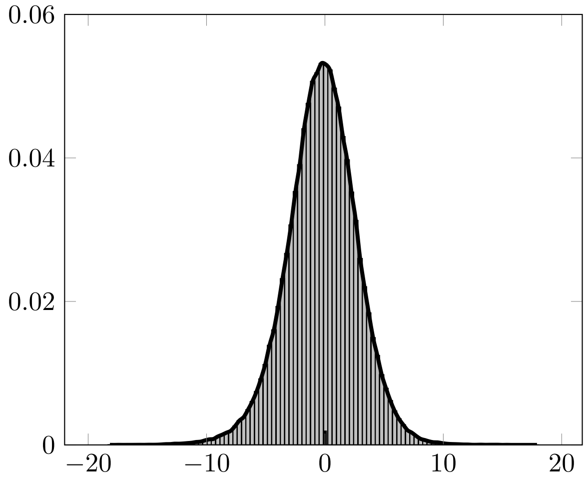


Figure A1. *Distribution of Word-Level Tone Scores. The tone scores across our corpus follow a roughly normal distribution with slight negative skewness (mean = -0.26, standard deviation = 2.95).*

words generate a negative similarity.

Finally, to capture the net distance from the two sets of seed words, the emotion score of each word in our corpus is calculated as follows:

$$s_i = \sum_{p \in P} \frac{w_i w_p}{||w_i||||w_p||} - \sum_{q \in Q} \frac{w_i w_q}{||w_i||||w_q||},$$

where P is the 100 positive seed words set and Q is the 100 negative seed words set. A positive score s_i indicates that w_i is closer to positive seed words in the vector space than to the negative ones.

Using this approach, we can assign a score to every word in our corpus of news articles. Therefore, we built an emotional word lexicon with approximately 400,000 words, which have been used at least 5 times in the corpus. Its distribution is close to the normal but slightly negatively skewed with a mean value of -0.26 and a standard deviation of 2.95. Figure A1 illustrates the distribution of the emotion scores of words.

In our study, the emotion score (or the extent of positivity/negativity) of each news article is an aggregate of words in its text. To generate the scores, the standard pre-processing procedures are routinely followed: We first obtain the stop words consisting of English stop words in nltk package along with punctuation marks and names. For each text, we eliminate the stop words and convert all capital letters to lower case letters, etc. In general, by utilizing the word lexicon, we calculate the article level emotional score by following the procedure below:

a. For each text, generate the sentences in the text and split those to obtain word list. Note that we do not drop duplicates words.

b. For each word in the word list, find the corresponding score in the word lexicon and add it to the text score.

c. On condition that a word has a internal negation right before it, such as "not satisfying", we assign the opposite emotion value of this word's to this phrase.

d. The score of text is the sum of word scores in the word list divided by number of words.

Three primary text scores are constructed by varying the word list in the texts. First, we construct word lists by using *all* the sentences in the texts. Second, we only include sentences that mention "China" or "Chinese." Third, we only include words whose emotion scores are far enough from the mean score of the lexicon, representing words with strong emotions, i.e., words whose scores are beyond one (or two) standard deviation(s) around the mean word score.

Article Level Validation

To validate our measure of tones at the article-level, we utilize human input as a validation. We randomly draw 100 articles from our sample, and then asked four trained assistants, all of whom are native English speakers, to independently evaluate tones of those articles, i.e., labelling them as "very very negative (-3)", "very negative (-2)""negative (-1)", "neutral (0)", "positive (1)", "very positive (2)"and "very very positive (3)". Note that in this validation process, evaluators were instructed to assess the overall tone of articles toward China, rather than specifically focusing on content about the Chinese government. We take the average of the individual scores as the average human rating for each article. We plot corresponding tone scores that are computed according to our algorithm against human ratings, as well as the fitted regression line in Figure A2. The estimated slope is 0.21 and it is highly significant, i.e., p-value is 0.005. There is a clear pattern whereby the computer algorithm and human rating largely agree on the underlying tones of the articles.

To present a more concrete impression of the results of the algorithm that we use to compute tones, we select three articles from the New York Times in our sample, which were rated as relatively neutral, very negative and very positive by our algorithm. Mindful of the fact that the median tone score of the New York Times articles in our main sample is -0.70; the most negative -2.3, and the most positive 2.0. Below are three corresponding examples from the section of Asia-Pacific of the New York Times. We only show sample sentences that mentioned China or Chinese.

An article with an around-median score is "Trump Embraces Foreign Aid to Counter

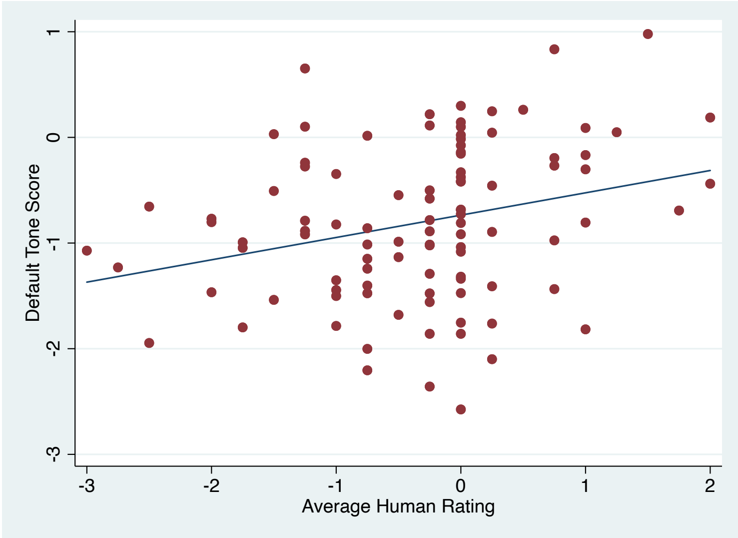


Figure A2. *Correlation between Algorithmic Tone Scores and Human Ratings. The scatter plot compares algorithmic tone scores (y-axis) with average human ratings (x-axis). The positive correlation (slope = 0.21, p = 0.005) demonstrates strong agreement between computational and human assessments of article tone.*

China's Global Influence (2018-10-14, score: -0.21)." Samples of sentences that mention "China or Chinese" are listed below:

> Mr. Trump seems to be learning that the projections of military power alone will not be enough to compete with *China*, he said.
>
> So much of our foreign policy now is focused on trying to check *China*, especially their nefarious activities.
>
> The key to its success, development officials said, is to create a new system that will carefully vet investments for maximum economic and political impact – and to ensure that projects don't fail as a result of corruption and mismanagement, a problem that has plagued *China*'s investments in Malaysia and elsewhere.
>
> A bigger question is whether it will do anything to reduce *China*'s global influence.

An article with very negative tone score is "Pneumonic Plague Is Diagnosed in China (2019-11-13, score: -2.28)." Samples of sentences that mention "China or Chinese" are listed below:

> On Tuesday, *Chinese* censors instructed online news aggregators in *China* to "block and control" online discussion related to news about the plague, according to a directive seen by The New York Times.

> Skeptical *Chinese* internet users have charged the government with being slow to disclose news about the disease, which is transmitted between humans and kills even faster than the more-common bubonic form.
>
> *China* has a history of covering up and being slow to announce infectious outbreaks, prompting many people to call for transparency this time.
>
> According to *China*'s health commission, six people have died in the country from the plague since 2014.

An article with very positive tone score is "Theater Director Returns to China With 'Liberating and Cool' Vision (2018-7-27, score: 1.58)." Samples of sentences that mention "China or Chinese" are listed below:

> In the way Chen Shi-Zheng imagines his theatrical adaptation of "The Orphan of Zhao," the production will bring out all the elements of the story that have appealed to *Chinese* audiences through the centuries, like the timeless themes of revenge and self-sacrifice.
>
> Over a recent dinner in New Haven, Mr. Chen and Audrey Li, his wife and business partner, talked with excitement about the chance for him to create a work for a *Chinese* audience again, playing the role of a cultural bridge as relations between the United States and *China* become more fraught over a variety of economic and security issues.
>
> After his formal arts education in *China*, he was invited to attend the Tisch School of the Arts at New York University as a graduate student, where he studied experimental theater from 1989 to 1991.

B. Additional Empirical Results and Discussions

B.1. Summary Statistics

Table A1. *Summary of Statistics*

	News			Opinions		
	Treatment mean (sd)	Control mean (sd)	Diff mean (se)	Treatment mean (sd)	Control mean (sd)	Diff mean (se)
	(1)	(2)	(3)	(4)	(5)	(6)
Default score	-0.73	-0.54	0.19	-0.76	-0.57	0.19
	(0.77)	(0.68)	(0.09)	(0.54)	(0.68)	(0.06)
China-Based score	-0.82	-0.59	0.23	-0.86	-0.60	0.25
	(0.87)	(0.80)	(0.14)	(0.64)	(0.77)	(0.06)
Score excluding 1 std	-1.59	-1.34	0.25	-1.51	-1.08	0.44
	(1.85)	(1.71)	(0.14)	(1.22)	(1.54)	(0.13)
Logged Wordcount	6.24	6.19	-0.05	6.65	6.13	-0.51
	(0.75)	(0.67)	(0.16)	(0.79)	(0.46)	(0.10)

Notes: This table presents means and standard deviations for treatment and control groups, separately for news and opinion articles. For each category, we report the difference in means between treatment and control groups. Standard deviations are shown in parentheses below means in columns (1), (2), (4), and (5). Standard errors in columns (3) and (6) are clustered at the press level to account for within-press correlation.

B.2. Placebo Tests with Entertainment News

Table A2. *Placebo Difference-in-Differences using Entertainment News Sample*

	Outcome Variable: Article-level Tone		
	(1)	(2)	(3)
Treated × Post	-0.022	-0.027	-0.015
	(0.038)	(0.042)	(0.039)
Treated	-0.286***	-0.222***	
	(0.049)	(0.029)	
Post	-0.210***	-0.218***	
	(0.036)	(0.039)	
Controls	No	Yes	Yes
Month FE	No	No	Yes
Press FE	No	No	Yes
Panel FE	No	Yes	Yes
R-Squared	0.055	0.150	0.182
N	96,794	96,794	96,794

Notes: This table presents difference-in-differences estimates for the sample of entertainment news articles. The dependent variable is the article-level tone score. Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. Column (1) shows the baseline specification without controls. Column (2) adds controls and panel fixed effects. Column (3) includes the full set of fixed effects (month and press) and controls. Controls include the logarithm of total word count and China-related terms ('China', 'Chinese', 'Taiwan'). Sample includes entertainment articles (which are excluded from our main analysis) with $\geq$ 3 China-related keywords. Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

B.3. Excluding One Outlet at a Time

Table A3. *Excluding One Outlet at One Time*

Excluding:	β	S.E.	p-value
Breitbart News	-0.153	0.0600	0.0180
Chicago Tribune	-0.153	0.0532	0.00888
The Dallas Morning News	-0.157	0.0530	0.00732
Huffpost	-0.151	0.0538	0.0103
New York Post	-0.164	0.0519	0.00458
The New York Times	-0.156	0.0555	0.0101
Star Tribune	-0.154	0.0533	0.00868
BBC	-0.144	0.0542	0.0144
The Boston Globe	-0.154	0.0540	0.00945
Daily Mail	-0.0838	0.0477	0.0926
Daily Mirror	-0.162	0.0519	0.00490
Financial Times	-0.154	0.0541	0.00926
The Guardian	-0.183	0.0483	0.00100
Independent	-0.137	0.0541	0.0188
Los Angeles Times	-0.152	0.0546	0.0110
Miami Herald	-0.158	0.0526	0.00666
NBC News	-0.161	0.0535	0.00649
Newsday	-0.154	0.0532	0.00843
San Francisco Chronicle	-0.164	0.0510	0.00402
Telegraph	-0.170	0.0501	0.00260
The Times	-0.156	0.0529	0.00730
USA Today	-0.155	0.0532	0.00806
The Washington Post	-0.176	0.0515	0.00253
The Wall Street Journal	-0.127	0.0514	0.0221

Notes: This table presents robustness checks where we re-estimate our main difference-in-differences specification (Equation (1)) while excluding one news outlet at a time. Each row shows the treatment effect (β), standard error, and p-value from a regression that excludes the specified outlet. The results remain statistically significant and economically similar across all specifications, suggesting that our findings are not driven by any single news outlet.

B.4. Driven by Trade war, Tiananmen or Hong Kong?

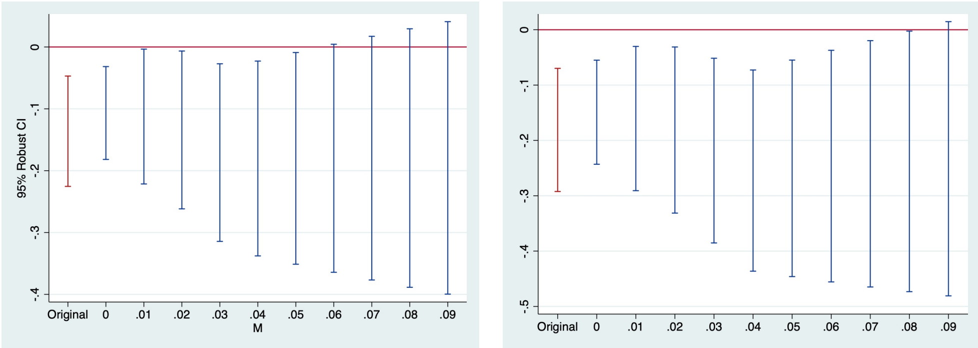
Do news articles mentioning trade war, TAM, or Hong Kong, the suspected triggers of this crackdown, drive our results? We conduct subsample analyses by excluding articles containing these keywords and reestimate Equation (1). Table A4 shows that excluding articles mentioning trade war (column 1), TAM (column 2), or Hong Kong (column 3) yields consistently significant negative effects (-0.169, -0.155, and -0.166 respectively, all significant at 1%), suggesting our main findings are robust and not driven by coverage of these events.

Table A4. *Excluding Suspected Triggers: Trade War, TAM and Hong Kong*

	Outcome Variable: Article-level Tone		
	Samples Excluding Articles that Mention:		
	Trade War	Tiananmen	HK
	(1)	(2)	(3)
Treated × Post	-0.169***	-0.155***	-0.166***
	(0.056)	(0.053)	(0.046)
Controls	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Press FE	Yes	Yes	Yes
Panel FE	Yes	Yes	Yes
R-Squared	0.235	0.197	0.185
N	38,892	46,705	36,277

Notes: This table presents difference-in-differences estimates excluding articles that mention specific events. Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. The dependent variable is the article-level tone score. All specifications include the full set of fixed effects (month, press, and panel) and controls. Column (1) excludes articles mentioning the US-China trade war. Column (2) excludes articles mentioning the Tiananmen Square incident (TAM). Column (3) excludes articles mentioning Hong Kong (HK). Controls include the logarithm of total word count and China-related terms ('China', 'Chinese', 'Taiwan'). Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

B.5. Honest Test for Parallel Trends



(a) Control Group I: All

(b) Control Group II: Always-blocked

Figure A3. *Sensitivity Tests*

Following Rambachan and Roth (2023), we conduct a sensitivity analysis based on smoothness restrictions, i.e. imposing that the slope of the difference in trends changes by no more than M between periods for our event study models.

Figure A3(a) shows the sensitivity test for our baseline DID result. The left of the figure shows the DID estimate in red, and then confidence intervals which get wider as we allow the slope to deviate by M. We can see that the blockage effect remains negative and statistically significant if the slope of the potential pre-existing trend differential between the treatment and the control group does not change ($M = 0$). Moreover, the breakdown value for a significant effect is $M = 0.06$, meaning that we can reject a null effect unless we are willing to allow for the linear extrapolation across consecutive periods to be off by more than 0.05. The result shows that our DID estimates are fairly robust.

Figure A3(b) shows the sensitivity test for our robustness test in which the control group only consists of the always-blocked media. We see that the breakdown value for a significant effect is $M = 0.09$. The result further shows the robustness of the DID estimates.

B.6. Testing for the Chilling Effect: Heterogenous Responses within the Control Group?

We test for potential chilling effects by relabeling always-blocked media as controls and never-blocked media as pseudo-treated, then estimating Equation (1). Table A5 shows results for default, China-based, and non-neutral tone scores (columns 1-3), defined in Section 3.1. The $T^{Pseudo} \times Post$ coefficients are consistently insignificant, suggesting no heterogeneous responses between groups. This absence of chilling effects aligns with the crackdown's targeting of media influence rather than content, further validating our control group selection.

Table A5. *Chilling Effects: Testing for Pre-emptive Self-Censorship*

	Outcome Variable:		
	Tone	China	Non-Neutral
	(1)	(2)	(3)
Treated$^{Pseudo}\times$ Post	-0.025	0.009	-0.112
	(0.059)	(0.084)	(0.073)
Controls	Yes	Yes	Yes
Press FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Panel FE	Yes	Yes	Yes
R-Squared	0.215	0.265	0.179
N	27,613	27,613	27,613

Notes: This table presents difference-in-differences estimates testing for potential chilling effects. We relabel always-blocked media as controls and never-blocked media as pseudo-treated for the control group that have not been blocked. Treated^{Pseudo} = 1 for media outlets never blocked; Post = 1 for months after January 2019. The dependent variables vary across columns: article-level tone score (column 1), tone scores of sentences that mention China (column 2), and tone scores of non-neutral tone indicators (which exclude the words with less emotive contents, within one std of the word tone distribution) (column 3). All specifications include the full set of fixed effects (month, press, and panel) and controls. Controls include the logarithm of total word count and China-related terms ('China', 'Chinese', 'Taiwan'). Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

B.7. Driven by Post-crackdown Events?

To address whether the harsher tone might reflect treated outlets' greater responsiveness to newsworthy events, particularly given potential unobservable differences between treatment and control groups, we conduct a robustness check. We remove articles covering the most salient post-blockage event - the COVID-19 pandemic - and reestimate the event study model using Equation (2). The results in Figure A4 show significantly negative coefficients both before the COVID-19 crisis and after April 2020, indicating our findings are not driven solely by pandemic-related coverage.

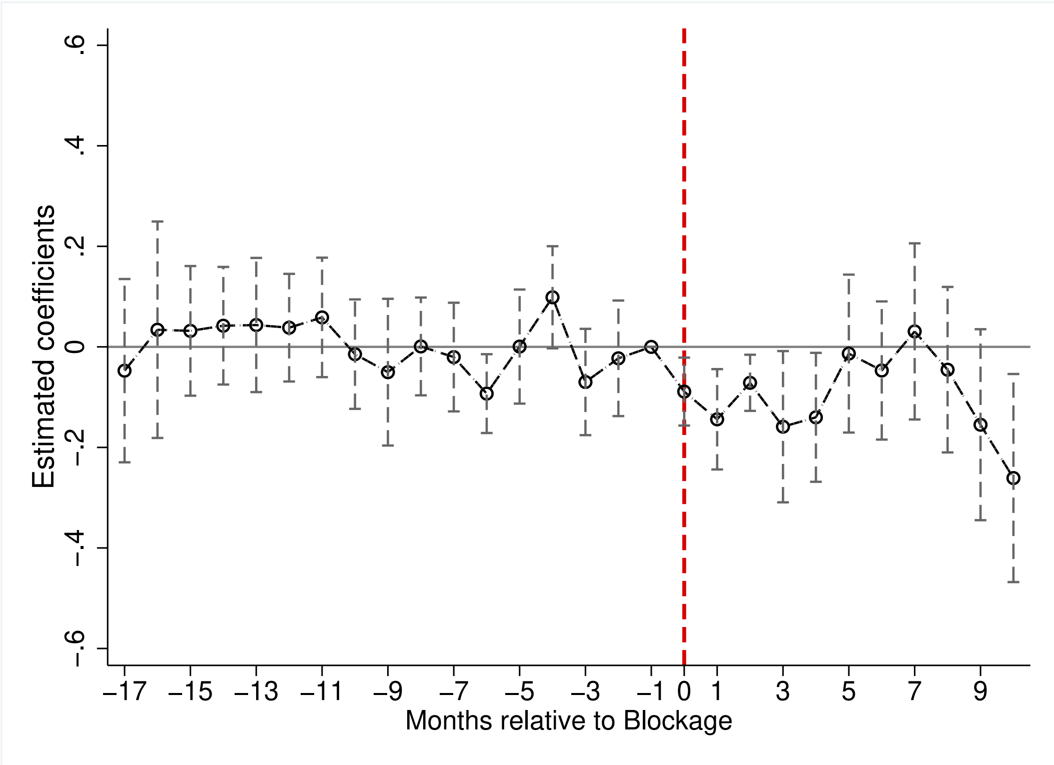


Figure A4. *Event study without COVID-19 articles. This figure illustrates coefficients and the associated confidence intervals estimated with the event study model and by using a subsample without COVID-19 articles. There is no difference in the preexisting trends between the treatment and control groups before the blockage. The timing of the divergence between the treatment and control groups coincides precisely with the crackdown. The month between before the crackdown is treated as the base period. Month*$_\tau$ *(where* $\tau = -17, \ldots, 10$*) represents dummy variables for the months from January 2018 to April 2020. In particular,* $\tau = -1$ *indicates the month of May 2019, at the end of which the crackdown occurred.*

B.8. Robustness Tests: Measurements and Samples

To examine whether the results are robust to the measure of tone, we reestimate Equation (1) with alternative measures discussed in section 3.2. Columns (1) and (2) of Table A6 report the results using the China-based scores and the nonneutral scores, respectively. The estimated blockage effects on the China-based scores and the nonneutral scores are -0.177 and -0.193, respectively.

Next, we test whether our results are robust to the choice of sample using articles with at least 5 or 1 keywords (discussed in section 3.1). The results in columns (3) and (4) of Table A6 yield estimates similar to those using the default news sample (column (2) of Table 3), suggesting robustness to sample choices.

In the main text, we use simple and transparent criteria for selecting China-related articles. Our default news sample contains articles with a minimum of three mentions of China-related keywords. To address potential type I and II errors in sample selection, we refine our sample by applying different criteria and filters on the default news sample with a minimum three mentions of China-related keywords. In Table A7, column (1) uses the articles that exclusively mention China-related terms without references to other countries in the headline; column (2) expands the default news sample by including articles from China news categories and with less than 3 mentions of China-related keywords; column (3) includes both the default news sample and the articles mentioning China-related terms in the headline but with less than 3 mentions of China-related keywords.

__Table A6.__ Robustness: Alternative Dependent Variables and Sample Restrictions

	Outcome Variable: China-Related	Outcome Variable: Non-Neutral	Outcome Variable: Tone	Outcome Variable: Tone
	Default Sample		Keywords ≥ 5	Keywords ≥ 1
	(1)	(2)	(3)	(4)
Treated × Post	-0.177**	-0.193**	-0.154**	-0.117**
	(0.067)	(0.084)	(0.063)	(0.049)
Controls	Yes	Yes	Yes	Yes
Press FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Panel FE	Yes	Yes	Yes	Yes
R-Squared	0.220	0.163	0.194	0.195
N	47,711	47,711	33,356	84,171

Notes: This table presents difference-in-differences estimates using alternative outcome measures and sample restrictions. Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. The dependent variables vary across columns: tone scores of sentences that mention China (column 1), tone scores of non-neutral tone indicators (which exclude the words with less emotive contents, within one std of the word tone distribution) (column 2), and article-level tone scores (columns 3-4). All specifications include the full set of fixed effects (month, press, and panel) and controls. Columns 1-2 use the default sample. Column 3 restricts to articles with 5 or more China-related keywords, while column 4 uses articles with 1 or more keywords. Controls include the logarithm of total word count and China-related terms ('China', 'Chinese', 'Taiwan'). Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

__Table A7.__ Robustness: Alternative Sample Construction Criteria

	Outcome Variable: Article-level Tone		
	Sample Restrictions		
	China-Only	China Category	China Headline
	(1)	(2)	(3)
Treated × Post	-0.161***	-0.155***	-0.153***
	(0.054)	(0.053)	(0.052)
Controls	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Press FE	Yes	Yes	Yes
Panel FE	Yes	Yes	Yes
R-Squared	0.199	0.197	0.196
N	37,290	47,791	48,066

Notes: This table presents difference-in-differences estimates using alternative sample construction methods for our main sample (articles with at least 3 mentions of China-related keywords). Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. The dependent variable is the article-level tone score. Columns vary by sample selection criteria: Column (1) excludes articles that mention other countries in their headlines; Column (2) adds all articles from news outlets' dedicated China sections; Column (3) adds articles that mention China in their headlines. All specifications include the same set of fixed effects and controls as in the baseline model. Controls include the logarithm of total word count and China-related terms ('China', 'Chinese', 'Taiwan'). Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

B.9. Summary Statistics for the Russia and Iran Samples

Table A8. *Summary of Statistics, Russia and Iran News Samples*

	Russia			Iran		
	Treatment mean (sd) (1)	Control mean (sd) (2)	Diff mean (se) (3)	Treatment mean (sd) (4)	Control mean (sd) (5)	Diff mean (se) (6)
Default score	-0.94	-1.00	-0.05	-1.43	-1.34	0.10
	(0.76)	(0.69)	(0.12)	(0.66)	(0.74)	(0.16)
Word count	617.21	466.62	-150.59	703.86	453.07	-250.79
	(775.58)	(356.38)	(68.38)	(1342.77)	(338.60)	(176.45)
Freq. Russia & Russian	8.74	7.97	-0.76	1.28	0.98	-0.29
	(7.99)	(6.83)	(0.53)	(4.60)	(3.09)	(0.176)
Freq. Iran & Iranian	0.10	0.10	-0.00	12.76	11.42	-1.34
	(0.37)	(0.36)	(0.03)	(11.96)	(10.28)	(1.29)
N	11105	13547	24652	6435	7497	13932

Notes: This table presents summary statistics for the sample of news articles covering Russia and Iran, comparing treatment and control groups. The Russia sample includes articles with at least 3 mentions of "Russia" or "Russian", while the Iran sample includes articles with at least 3 mentions of "Iran" or "Iranian". Treatment group consists of media outlets that were blocked, while control group includes those that were never blocked. Article tone score, with more negative values indicating more negative sentiment, varies slightly across treatment and control groups. Word count represents the total number of words per article. Frequency measures count the number of country-specific keyword mentions in each article. The differences (Diff) columns report the mean differences between treatment and control groups, with standard errors clustered at the press level shown in parentheses. Standard deviations (sd) are shown in parentheses below the means in columns 1-2 and 4-5.

B.10. Additional Tests for Mechanisms

Table A9. *Heterogeneous Effects: The Role of Journalistic Resources?*

	Outcome Variable: Article-level Tone	
	(1) Chinese Platform	(2) Chinese Influence
Treated × Post	-0.178***	-0.190***
	(0.058)	(0.058)
Post × Chinese Platform	0.071	
	(0.060)	
Treated × Post × Chinese Platform	0.132*	
	(0.064)	
Post × High Baidu		0.072
		(0.060)
Treated × Post × High Baidu		0.139**
		(0.062)
Controls	Yes	Yes
Month FE	Yes	Yes
Press FE	Yes	Yes
Panel FE	Yes	Yes
R-Squared	0.199	0.199
N	47,711	47,711

Notes: This table examines the heterogeneous effects of the crackdown event across outlets with Chinese online platforms and higher influences and those without. The dependent variable is the article-level tone score. Column (1) investigates heterogeneous effects based on whether the outlet has Chinese platforms, while column (2) explores the heterogeneous effects based on Chinese influence. Treated = 1 for media outlets blocked in the May of 2019; Post = 1 for the post-treatment period. Chinese Platform indicates whether the outlet has Chinese platform publishing and translating their news content into Chinese, and High Baidu indicates high Chinese influence measured by Baidu search index of this outlet above the median in this sample. All specifications include press, month, and panel fixed effects. Controls include article-level characteristics. The results suggest that while the treatment decreased tone by 0.178 standard deviations, this effect is partially mitigated (by 0.132) for outlets with Chinese platform and (by 0.139) for those with high Chinese influence. Standard errors in parentheses are clustered at the press level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

Table A10. *COVID-19 Shock and News Coverage: The Role of Journalistic Resources*

	Outcome Variable: Article-level Tone	
	(1)	(2)
Post COVID-19 × Chinese Platform	0.103**	0.137***
	(0.049)	(0.045)
Post COVID-19	-0.269***	
	(0.039)	
Chinese Platform	0.068	
	(0.102)	
Controls	Yes	Yes
Month FE	No	Yes
Press FE	No	Yes
Panel FE	Yes	Yes
R-Squared	0.150	0.229
N	21,364	21,364

Notes: This table examines how media outlets' tone responded to the COVID-19 shock, using whether they maintain Chinese-language platforms to proxy the adjustment of journalistic resources. The dependent variable is the article-level tone score, which measures the sentiment of news coverage, with more negative values indicating more negative sentiment. Post COVID-19 = 1 for months after January 2020; Chinese Platform = 1 for media outlets that maintain Chinese-language website versions. Column 1 presents the baseline specification with panel fixed effects only, while column 2 includes the full set of fixed effects (month, press, and panel). Controls include the logarithm of total word count and China-related terms ('China', 'Chinese', 'Taiwan'). The interaction term captures the differential effect of the COVID-19 shock on outlets with Chinese platforms. Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

Table A11. *A Composition Change in Audience's Attention*

	Outcome Variable: Article-level Tone	
	(1)	(2)
Treated × Post	-0.162***	-0.149***
	(0.050)	(0.044)
ln(Baidu index)	-0.074	-0.104**
	(0.050)	(0.039)
ln(Google index)	-0.019	-0.019
	(0.015)	(0.016)
ln(Baidu index) × *Post*		0.048*
		(0.025)
ln(Google index) × *Post*		0.003
		(0.020)
Controls	Yes	Yes
Month FE	Yes	Yes
Press FE	Yes	Yes
Panel FE	Yes	Yes
R-Squared	0.197	0.199
N	47,711	47,711

Notes: This table examines how changes in audience composition affect news coverage. The dependent variable is the article-level tone score, which measures the sentiment of news coverage, with more negative values indicating more negative tone. Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. The Baidu index and Google index measure search interest from mainland Chinese and international audiences, respectively. Column 1 presents the baseline specification with time-varying Baidu and Google indices for each newspaper as additional controls, while column 2 adds interaction terms between these search indices and Post. All specifications include the full set of fixed effects (month, press, and panel). Controls include the logarithm of total word count and China-related terms ('China', 'Chinese', 'Taiwan'). The interaction terms capture how changes in audience attention affect coverage tone differently before and after the blocking. Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

C. Topic Modeling

To estimate the Latent Dirichlet Allocation (LDA) model, we pre-processed our news corpus following standard practices. We converted words to lowercase, removed stop words (e.g., "a", "an", "the"), punctuation, numbers, white space, and URLs, while preserving dashes within words. After stemming words and retaining only terms occurring at least five times, our vocabulary size became 40,466.

We chose $K = 12$ topics for the LDA model (justified in the main text) and fitted it using Gibbs sampling, following Blei, Ng, and Jordan (2003)'s algorithm implemented in R's *topic models* package. The estimation stabilized after 1,000 iterations, with similar results for $K = 11$ or $K = 13$.

From the estimation outputs, we focused on the most frequently used words in each topic and the distribution of documents over topics. The resulting topics were interpretable and intuitive, corresponding to identifiable China-related events during the data period.

Tables A12 and A13 show the topics' highest-probability words, while Figure A5 presents word clouds for topics discussed in the main text, with word sizes proportional to their topic probabilities.

Table A12. *Top Word Lists*

Topic 1 Growth	Topic 2 Trade	Topic 3 Companies	Topic 4 US Affairs	Topic 5 COVID/Report	Topic 6 Human Rights
market	china	compani	trump	peopl	china
year	trade	china	presid	coronavirus	chines
economi	trump	huawei	say	will	govern
growth	tariff	year	will	health	report
bank	chines	technolog	hous	countri	offici
stock	american	chines	time	test	beij
trade	presid	busi	think	report	parti
price	import	product	want	govern	media
rate	state	use	like	death	state
per	administr	billion	one	day	foreign
month	deal	firm	get	new	communist
global	unit	will	know	case	author
econom	good	new	democrat	number	nation
investor	will	million	just	week	right
will	billion	sale	can	state	countri
cent	war	execut	may	home	intern
expect	beij	industri	now	virus	post
percent	product	appl	peopl	outbreak	investig
last	countri	invest	state	now	inform
point	econom	oper	deal	spread	use
also	negoti	accord	new	need	accord
sinc	offici	includ	polit	work	alleg
financi	talk	make	elect	offici	also
cut	export	also	right	public	social
week	washington	last	call	also	univers
fell	impos	develop	white	itali	secur
index	agreement	govern	make	infect	law
rise	two	plan	back	can	accus
quarter	steel	network	vote	hospit	one
china	economi	servic	thing	close	time
share	percent	market	former	travel	year
report	hous	group	work	covid-	human
increas	make	secur	american	includ	call
drop	busi	tech	ask	pandem	public
gain	industri	can	look	announc	polit
invest	polici	share	see	mask	peopl
fall	includ	like	way	get	claim
oil	also	data	leader	measur	arrest
higher	global	one	first	worker	say
rose	white	custom	campaign	one	comment

Table A13. *Top Word Lists (Cont'd)*

Topic 7 NK/TW/Russia	Topic 8 Social Issues	Topic 9 HK protests	Topic 10 Miscellaneous	Topic 11 COVID/Travel	Topic 12 COVID/Outbreak
china	year	hong	news	new	virus
north	one	kong	pictur	case	peopl
korea	famili	protest	world	coronavirus	coronavirus
countri	show	polic	protest	china	infect
south	peopl	china	fire	dead	can
militari	man	govern	ralli	travel	patient
state	home	citi	take	wuhan	china
kim	time	peopl	near	flight	case
korean	citi	offic	day	soar	spread
taiwan	children	demonstr	minist	flu-lik	health
presid	live	beij	nation	australia	test
unit	day	chines	elect	passeng	outbreak
will	women	bill	burn	hospit	diseas
nation	imag	mainland	offici	confirm	wuhan
russia	video	law	area	airport	one
leader	build	one	polic	ship	two
iran	first	forc	part	arriv	will
meet	told	extradit	member	man	may
nuclear	two	lam	support	peopl	first
offici	around	month	hold	quarantin	ill
also	get	use	prime	protect	expert
sanction	last	support	block	australian	around
region	just	movement	opposit	airlin	scientist
beij	also	call	california	chines	drug
secur	back	street	public	two	develop
summit	school	gas	woman	health	say
two	found	march	coronavirus	citi	caus
intern	film	polit	mask	return	studi
forc	new	violenc	covid-	woman	work
chines	dog	fire	forc	staff	symptom
report	life	carri	san	wear	use
war	can	demand	call	mask	report
visit	million	leader	peopl	januari	anim
power	mother	legisl	mark	intern	world
year	left	two	seen	-year-old	human
sea	woman	right	face	south	medic
island	polic	sunday	along	virus	time
foreign	accord	arrest	home	cruis	research
india	three	tear	thousand	outbreak	also
japan	moment	continu	west	plane	offici

Table A14. *Economic Topics: Intensive Margin*

	Outcome Variable: Article-level Tone				
	Topic 1 Market (1)	Topic 2 Trade (2)	Topic 3 Companies (3)	Topic 4 US (4)	Topic 5 COVID Report (5)
Treated × Post	0.061	0.041	0.066	-0.030	-0.067
	(0.052)	(0.040)	(0.055)	(0.051)	(0.053)
Controls	Yes	Yes	Yes	Yes	Yes
Press FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Panel FE	Yes	Yes	Yes	Yes	Yes
R-Squared	0.162	0.193	0.178	0.192	0.262
N	11,928	11,928	11,928	11,928	11,928

Notes: This table examines how the blocking event affects article tone across different topics. The dependent variable is the article-level tone score, which measures the sentiment of news coverage, with more negative values indicating more negative tone. We construct topic-specific subsamples by selecting articles in the top quartile of topic probability based on LDA topic modeling and re-estimate the baseline difference-in-differences specification. Each column represents a different topic subsample: market-related (1), trade-related (2), company-related (3), US-related (4), and COVID-related (5). Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. All specifications include the full set of fixed effects (month, press, and panel). Controls include the logarithm of total word count and China-related terms ('China', 'Chinese', 'Taiwan'). Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

Table A15. *Politically Sensitive Topics: Intensive Margin*

	Outcome Variable: Article-level Tone						
	Human rights (1)	NK/TW/Russia (2)	Social (3)	HK (4)	Miscellaneous (5)	COVID Travel (6)	COVID Outbreak (7)
Treated × Post	-0.116***	-0.127***	-0.123***	-0.108**	-0.110**	-0.137**	-0.155**
	(0.041)	(0.040)	(0.039)	(0.044)	(0.041)	(0.050)	(0.056)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Press FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.212	0.150	0.160	0.217	0.228	0.238	0.278
N	11,928	11,928	11,928	11,928	11,928	11,928	11,929

Notes: This table examines how the crackdown event affects article tone across different politically sensitive topics. The dependent variable is the article-level tone score, which measures the sentiment of news coverage, with more negative values indicating more negative tone. We construct topic-specific subsamples by selecting articles in the top quartile of topic probability based on LDA topic modeling and re-estimate the baseline difference-in-differences specification. Each column represents a different topic subsample: human rights-related (1), North Korea/Taiwan/Russia-related (2), social issues (3), Hong Kong-related (4), miscellaneous (5), COVID travel restrictions (6), and COVID outbreak news (7). Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. All specifications include the full set of fixed effects (month, press, and panel). Controls include the logarithm of total word count and China-related terms ('China', 'Chinese', 'Taiwan'). Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

Table A16. *Excluding One Topic at a Time*

Excluded Topic	Coefficient	Std. Error	p-value
Market	-0.229	0.042	0.000
Trade	-0.225	0.046	0.000
Chinese Companies	-0.197	0.048	0.000
US Affairs	-0.210	0.058	0.001
Human Rights	-0.189	0.065	0.008
NK/Russia affairs	-0.141	0.077	0.078
Social	-0.092	0.043	0.043
Hong Kong Protest	-0.155	0.067	0.030
Miscellaneous	-0.166	0.054	0.005
COVID	-0.105	0.037	0.009

Notes: This table presents regression results after excluding specific topics from the analysis. Each row represents a separate regression where the indicated topic is excluded. The coefficients represents the DID estimates for the blockage effects, with their corresponding standard errors and p-values. All standard errors are clustered at the outlet level.

__Table A17.__ Summary statistics: Weekly number of articles.

		Treatment		Control	
Topic Number	Topic Name	mean	sd	mean	sd
1	Market and Growth	5.03	5.71	4.11	6.33
2	Trade	5.64	5.04	3.90	5.39
3	Companies	5.21	5.04	4.05	5.86
4	US Affairs	8.11	7.65	3.04	4.52
5	COVID Report	8.57	22.80	2.88	5.84
6	Human Rights	8.39	8.31	2.94	3.54
7	N.Korea, Taiwan, Russia	6.94	6.07	3.44	4.33
8	Social Issues	9.04	14.25	2.71	3.68
9	Hong Kong Protests	7.22	7.95	3.34	3.96
10	Miscellaneous	6.31	7.14	3.67	5.36
11	COVID/Travel	9.12	22.13	2.68	4.01
12	COVID/Outbreak2	9.59	26.00	2.52	4.38

Notes: This table presents summary statistics for the weekly number of articles across different topics, comparing treatment and control groups. Topics are identified through LDA topic modeling. The first four topics (1-4) are primarily economic in nature, while the remaining topics (6-9) are more politically sensitive. Treatment group consists of media outlets that were blocked in May 2019, while the control group includes never blocked outlets and outlets that were blocked earlier than 2018. Mean represents the across outlet average weekly number of articles per topic, and sd represents the standard deviation.

Table A18. *Economic Topics: Extensive Margin*

	Outcome Variable: Weekly Number of Articles				
	Topic 1 Growth (1)	Topic 2 Trade (2)	Topic 3 Companies (3)	Topic 4 US affairs (4)	Topic 5 COVID Report (5)
Treated $\times$ Post	1.139	0.509	0.760	1.913***	0.697
	(0.676)	(0.780)	(0.583)	(0.561)	(1.239)
Press FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
R-Squared	0.804	0.708	0.821	0.784	0.783
N	2,742	2,742	2,742	2,742	2,742

Notes: This table examines how the crackdown event affects the volume of coverage across different topics. The dependent variable is the weekly number of articles for each topic, identified through LDA topic modeling. Each column represents a different topic: growth-related (1), trade-related (2), company-related (3), US affairs-related (4), and COVID report-related (5). Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. All specifications include press and month fixed effects. Controls include the monthly count of articles in all topics combined for each outlet to capture the size of the outlet. The results suggest that the crackdown event had a significant positive effect only on the volume of US affairs coverage, with an increase of approximately 1.9 articles per week. Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

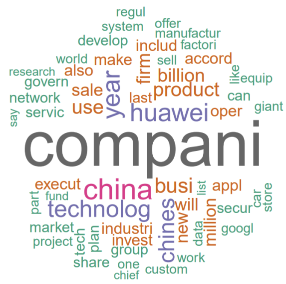
Table A19. *Politically Sensitive Topics: Extensive Margin*

	Outcome Variable: Weekly Number of Articles						
	Human rights (1)	NK/Taiwan/Russia (2)	Social (3)	HK (4)	Miscellaneous (5)	COVID Travel (6)	COVID Outbreak (7)
Treated × Post	1.807^{*}	-0.357	-0.912	3.868^{***}	-3.157^{**}	0.081	0.367
	(0.966)	(0.585)	(0.768)	(1.068)	(1.189)	(0.945)	(1.106)
Press FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.787	0.785	0.930	0.747	0.697	0.787	0.890
N	2,742	2,742	2,742	2,742	2,742	2,742	2,742

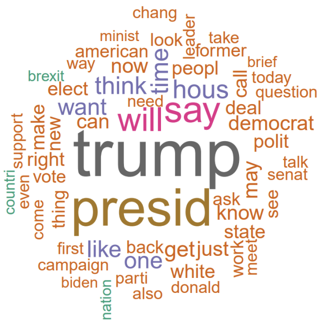
Notes: This table examines how the crackdown event affects the volume of coverage across different politically sensitive topics. The dependent variable is the weekly number of articles for each topic, identified through LDA topic modeling. Each column represents a different topic: human rights-related (1), North Korea/Taiwan/Russia-related (2), social issues (3), Hong Kong-related (4), miscellaneous political topics (5), COVID travel restrictions (6), and COVID outbreak news (7). Treated = 1 for media outlets blocked in May 2019; Post = 1 for months after May 2019. All specifications include press and month fixed effects. Controls include the monthly count of articles in all topics combined for each outlet to capture the size of the outlet. The results show significant effects for several topics: increases in Hong Kong coverage (3.9 articles/week), human rights coverage (1.8 articles/week), and a decrease in miscellaneous coverage (-3.2 articles/week). Standard errors in parentheses are clustered at the media outlet level. Statistical significance: * p<0.1, ** p<0.05, *** p<0.01.

(a) Topic 1

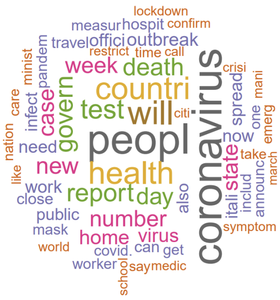
(b) Topic 2



(c) Topic 3



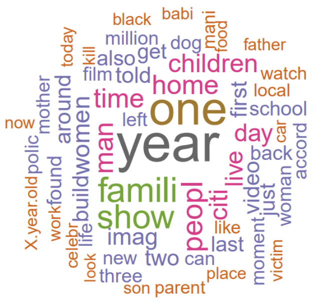
(d) Topic 4



(e) Topic 5

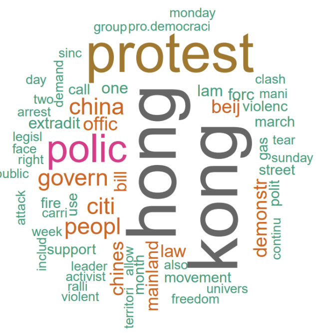
(f) Topic 6

(g) Topic 7

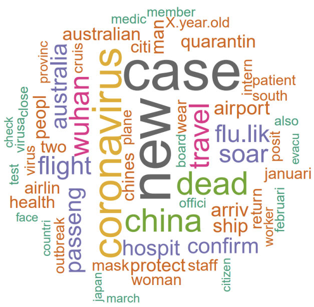


(h) Topic 8

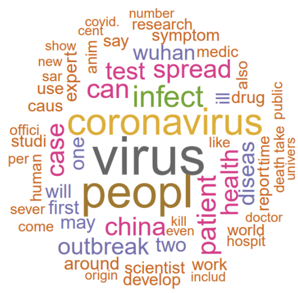
(i) Topic 9



(j) Topic 10



(k) Topic 11



(l) Topic 12

Figure A5. *Word Clouds: Word Size Proportional to Topic-Word Probabilities from LDA Topic Modeling*