The War of Ideas: Institutions and Global Media Bias *

Sibo Liu[†], Alexey Makarin[‡], Jinfeng Wu[§], Dong Zhang[¶]

February 2026

Abstract

We present the first systematic analysis of how countries cover each other in the news, using a newly constructed dataset of 55,240 media sources from 176 countries between 2015 and 2022, drawing upon a corpus of 1.4 billion online news articles. We document a pattern of global media bias: countries with greater differences in democracy levels systematically portray each other more negatively, shaping both the sentiment and thematic content of coverage. This bias is especially pronounced when less democratic countries report on liberal democracies, when state-controlled media is involved, or when the reporting country experiences economic downturns or political turmoil. We corroborate these findings exploiting episodes of democratic erosion in a difference-in-differences design. Using similarities in how countries report on *third* countries, we construct and introduce a novel granular measure of global media alliances. Finally, we show that cross-country media sentiment predicts subsequent public attitudes.

Keywords: media bias, democracy, information control, geopolitics, soft power

JEL Codes: P00, F51, D74, L82

*We are grateful to Daron Acemoglu, Charles Angelucci, Elliott Ash, Giulia Caprini, Annalí Casanueva Artís, Ruben Durante, Ruben Enikolopov, Germain Gauthier, Pierre-Guillaume Méon, Andrea Prat, Jesse Schreger, Jesse Shapiro, David Strömberg, David Yanagizawa-Drott, and David Yang for insightful comments. We also thank participants at the CEPR Workshop on Media, Technology, Politics, and Society; the CESifo Venice Workshop on Digital Platforms; the 2025 APSA Annual Meeting; the Zurich Workshop in AI+Economics; and seminars at Hong Kong Polytechnic University, the ifo Big Data Seminar, and Lingnan University.

[†]School of Accounting and Finance, Hong Kong Polytechnic University
(email: sibo.liu@polyu.edu.hk)

[‡]MIT Sloan School of Management, CEPR, and CESifo (email: makarin@mit.edu)

[§]Division of Social Science, HKUST (email: jwudb@connect.ust.hk)

[¶]Division of Social Science, HKUST (email: dongzhang@ust.hk)

"Advances in information technology, satellite broadcasting, and the Internet will enable Western media networks to saturate our domestic audience with their reports and views. Countries that try to block the use of IT will lose. We have to learn to manage this relentless flood of information so that the Singapore government's point of view is not smothered by the foreign media."

—Lee Kuan Yew, From Third World to First (2000: p.196)

1 Introduction

On February 24, 2022, news of Russia's military incursion into Ukraine quickly spread across the globe, but the framing of these events often differed sharply. In the United States, Fox News led with the title *"Russia invades Ukraine in largest European attack since WWII,"* conveying a sense of historical gravity and casting Russia as the clear aggressor (Fox News, 2022). Across the Atlantic, *The Guardian* similarly ran the front page headline *"Putin invades"* (The Guardian, 2022). By contrast, China's state-run *People's Daily* reported that *"Russia-Ukraine conflicts intensify, sending shock waves across world,"* a more neutral headline that avoided explicit language of aggression (People's Daily, 2022). Similar asymmetries arise elsewhere: for example, coverage by U.S. outlets of the war in Yemen, waged by a Saudi-led coalition with U.S. support, tended to employ more-muted language with little direct responsibility attribution (Bachman and Brito Ruiz, 2024). Far from being mere editorial differences, these contrasting portrayals reflect deeper forces that shape how each country's media outlets depict foreign events—a phenomenon we term *global media bias*.

Although objectivity, balance, and truthfulness are widely viewed as the pillars of journalistic ethics, a substantial body of work in economics and political science shows that media outlets often deviate from these ideals in practice. Researchers have documented pervasive ideological or partisan leanings across various democratic contexts, indicating that media bias is neither isolated nor incidental (Groseclose and Milyo, 2005; Gentzkow and Shapiro, 2010; Larcinese, Puglisi, and Snyder, 2011; Durante and Knight, 2012; Martin and Yurukoglu, 2017; Ash et al., 2021; Caprini, 2023). Yet whether and how these biases extend beyond national borders to systematically influence cross-country reporting remains an open question, with potentially important implications for the formation of international perceptions and attitudes.

An important dimension along which media coverage may systematically differ is political-regime distance between countries. When regime types diverge, media outlets—particularly those subject to state influence—may face strategic incentives to portray foreign countries more negatively, shaping domestic beliefs while also targeting audiences abroad. Such incentives can be amplified by concerns over political legitimacy or geopolitical competition. Consistent with this view, policy analysts suggest that China has mobilized its state media apparatus to wage a global public opinion "war of ideas" against the United States and the West, promoting China's image and regime while seeking to discredit and delegitimize its rivals (Holz, 2024).

This paper provides the first systematic analysis of how countries cover one another in the news. We organize the analysis around four main questions. First, do systematic patterns of media bias emerge in cross-country reporting, particularly between countries with different political regimes? Second, what factors shape the magnitude and direction of such bias? Third, can similarities in how countries report on third countries be leveraged to shed light on the geopolitical media landscape? Fourth, does biased international news coverage have measurable relevance for political attitudes and countries' public image?

Answering these questions poses several major empirical challenges. First, there is a lack of standardized datasets that systematically track how media outlets across countries report on foreign countries, making it difficult to study international patterns of media coverage at scale. Second, media systems differ widely across countries in their institutional structure, language, and cultural context, complicating cross-country comparisons and necessitating a standardized and scalable measure of media bias. Third, isolating the determinants of media coverage is inherently difficult, as it likely reflects a combination of outlet-specific preferences, political and cultural developments in both the source and target countries, and the topics being reported.

To address these challenges, we draw on an extensive corpus of 1.4 billion online news articles to construct a comprehensive dataset of global media coverage, spanning 2015 to 2022 and comprising 55,240 distinct sources across 176 countries. These data are sourced from the Global Knowledge Graph (GKG) 2.0 of the GDELT Project, which provides real-time translation as well as emotion and theme detection of online news in 65 languages. Importantly, GKG 2.0 reports a tone score for each article that captures the overall sentiment of its content. Following Puglisi and Snyder (2015), we interpret systematic variation in tone as a measure of media bias, and we define coverage as biased when an outlet *systematically* portrays a foreign country in a more favorable or unfavorable light. We aggregate the article-level data to the outlet–target-country–theme–month level.

Building on this dataset, we examine whether international media coverage exhibits systematic biases along ideological and geopolitical lines. Our baseline specification is a gravity-style model in which we regress the average tone of a media outlet toward each foreign country, within a given theme and month, on the difference in political regimes (hereafter, *institutional distance*) between the two countries in the previous year. We measure *institutional distance* as the absolute difference in liberal democracy scores from the Varieties of Democracy (V-Dem) project (Coppedge et al., 2023), which captures differences in regime type and can also be interpreted as a proxy for broader geopolitical divergence.[1]

The granularity of our data allows us to include a rich set of fixed effects that account for a wide range of time-invariant characteristics and time-varying shocks, in the spirit of the gravity models in international economics (Anderson and Van Wincoop, 2003; Larch and Yotov, 2024). Our preferred specification includes outlet-by-theme-by-year and target-country-by-theme-by-year fixed effects, which flexibly absorb general tone shifts driven by domestic or foreign events, changes in outlet-level editorial policy that are independent of the specific foreign country being covered, and global thematic trends. Further, this demanding structure ensures that our coefficient of interest is estimated from year-to-year changes in *relative* institutional distance between countries, net of any uniform shifts in tone associated with the *absolute* democracy levels of the source or target country within a given theme-year.

Our baseline findings document a strong and robust negative relationship between the tone of a country's media coverage of another country and the institutional distance between them, consistent with the presence of systematic, ideologically driven global media bias. Specifically, when the contrast in levels of democracy between two countries is greater, media outlets in the source country tend to portray the target country in a more negative light. This relationship holds when we control for other

[1] Our findings are robust to alternative democracy scores, such as the Polity Index. We also examined media bias by indicators of geopolitical alignment, such as ideal point distance in United Nations General Assembly (UNGA) voting (Voeten, 2000, 2013; Bailey, Strezhnev, and Voeten, 2017). While UN voting distance is also negatively associated with media tone, its explanatory power disappears once we control for institutional distance. We therefore focus on institutional distance as the primary measure of analysis, as it appears to be more relevant for our context.

potential determinants of relationships between countries—such as common language (Melitz and Toubal, 2014), religious distance (Dow and Karunaratna, 2006), genetic distance (Spolaore and Wacziarg, 2009), geographic distance, and international trade flows (Kleinman, Liu, and Redding, 2024)—and also in an alternative specification when we include country-pair fixed effects that absorb all time-invariant bilateral heterogeneity. The results are also robust to a range of sensitivity checks, including the use of alternative tone measures and alternative samples.

To address concerns that the GKG 2.0 article repository may be overly broad and include topics less relevant to our analysis, we restrict our attention to a focused subset of news articles that mention national leaders—a domain widely recognized as geopolitically salient (e.g., Saich, 2014). In this refined sample, the relationship between institutional distance and negative tone becomes substantially stronger, suggesting that our baseline estimates may, if anything, understate the extent of ideologically driven global media bias.

Further, we show that institutional distance shapes not only tone but also topic selection in international news coverage. We find that media outlets devote more coverage to generally more negative themes when reporting on institutionally distant countries. This pattern suggests that ideological and political differences influence not only *how* foreign countries are portrayed but also *which* aspects of their lives and behavior are emphasized, thereby contributing to systematically biased narratives.

Next, we explore the role of government information control in shaping international reporting (e.g., Guriev and Treisman, 2019, 2020b) by documenting several patterns of heterogeneity. First, the relationship between institutional distance and media bias is weaker when the source country is a liberal democracy, but significantly stronger when the target country is a liberal democracy. Second, media bias in cross-national coverage is significantly more pronounced when the source country exhibits higher levels of traditional and internet censorship. Third, media bias is stronger in countries with greater state ownership of media outlets, including print and broadcast media. Fourth, media bias is more pronounced during economic distress in the source country, such as downturns or surging inflation. Fifth, media bias is stronger during episodes of domestic political turmoil in less democratic source countries. These patterns are consistent with an information-control interpretation of our baseline results: governments, particularly authoritarian ones, face stronger incentives to manipulate international narratives during times of domestic stress in order to manage public perceptions and maintain political stability (e.g., Martinez, 2022).

To complement our baseline analysis and assess causality more directly, we exploit episodes of democratic erosion in the source country using a difference-in-differences (DID) design. Our specification includes an interaction between the level of democracy in the target country and an indicator for democratic backsliding (or autocratization) in the reporting country, measured as a shift in a country's V-Dem regime classification toward less democratic categories. We find that, following democratic backsliding, media coverage from the affected country becomes significantly more negative toward highly democratic countries relative to less democratic ones. We find no evidence of pretrends, and the results are robust to using the estimator proposed by Sun and Abraham (2021), which addresses potential biases from negative weighting in staggered treatment designs. These findings are consistent with an interpretation in which democratic backsliding coincides with systematic changes in state-aligned narratives toward more-negative portrayals of liberal democracies, perhaps combined with heightened political control or pressure on domestic media.

In addition to documenting systematic biases in international media coverage, we introduce a novel approach to measuring geopolitical and ideological alignment between countries based on the co-movement of their media narratives. This method builds on the idea that alliances in the modern world are not only forged through formal treaties or voting patterns but also revealed in how countries frame and prioritize global events. We construct a correlation-based measure of alignment between country pairs by assessing the extent to which their media outlets report on third countries in a similar way across a range of thematic dimensions. This yields a granular, high-frequency, and scalable indicator of ideological and political proximity—an empirical construct that remains rare in the existing literature (Mohr and Trebesch, 2024, p. 19). The resulting network reveals clear ideological clustering: liberal democracies tend to group around the United States and Great Britain, while less democratic regimes form separate, smaller hubs, with Russia and China on the periphery. We show that this measure is positively related to standard indicators of geopolitical alignment. Finally, we document a decline in inter-regime connectivity over time, consistent with intensifying geopolitical tensions between countries with differing regime types.

Finally, we examine the relevance of global media bias for public opinion. Using Gallup World Poll data, we find that more-favorable sentiment in a country's media coverage of another country predicts a subsequent increase in public favorability toward the latter. While not necessarily causal, this pattern is consistent with the view that international media narratives can shape broader shifts in public opinion and that governments may seek to leverage such narratives as a channel of soft power.

This paper contributes to the literature in several ways. First, it provides the first systematic analysis of the global media landscape. In particular, we document systematic biases in how countries' media outlets cover other nations, depending on differences in political regimes and geopolitical alignment. In doing so, the paper advances our understanding of the interplay between media and politics (for reviews, see Prat and Strömberg, 2013; Strömberg, 2015; Zhuravskaya, Petrova, and Enikolopov, 2020). Most existing empirical work focuses on ideological or partisan bias in domestic contexts (e.g., Groseclose and Milyo, 2005; Gentzkow and Shapiro, 2010; Larcinese, Puglisi, and Snyder, 2011; Durante and Knight, 2012; Martin and Yurukoglu, 2017; Qin, Strömberg, and Wu, 2018; Ash et al., 2021; Caprini, 2023). Some prior work has examined specific cases; for example, Chen and Han (2022) show that U.K.- and U.S.-based outlets adopted a more negative tone toward China after being blocked by the Chinese government.[2] In contrast, our analysis adopts a comprehensive, comparative perspective across a broad range of countries and thematic domains, allowing us to document systematic patterns in international media bias and characterize the structure of the entire available global media landscape.

Second, our findings closely align with the literature on information control in nondemocratic regimes (e.g., Egorov, Guriev, and Sonin, 2009; Gehlbach and Sonin, 2014; Chen and Yang, 2019; Guriev and Treisman, 2019, 2020b; Martinez, 2022; Simonov and Rao, 2022).[3] Recognizing the destabilizing consequences associated with independent media and social media (Enikolopov, Petrova, and Zhuravskaya, 2011; Acemoglu, Hassan, and Tahoun, 2018; Enikolopov, Makarin, and Petrova, 2020; Guriev, Melnikov, and Zhuravskaya, 2021), authoritarian governments often resort to censorship measures to restrict the free flow of information (King, Pan, and Roberts,

[2]See also Qian and Yanagizawa (2009) for evidence of strategic bias in official U.S. human rights reports during the Cold War; Baum and Zhukov (2015, 2019) and Shaver et al. (2022) on biases in media-based conflict datasets; Hatte, Madinier, and Zhuravskaya (2021) on how social media availability shapes foreign reporting on the Israel-Palestine conflict; Durante and Zhuravskaya (2018) on how international media coverage influences strategic behavior in conflict settings; and Fetzer and Garg (2025) on empathy-driven international coverage of natural disasters specifically.

[3]We recognize that in both established and weak democracies, incumbent political parties and governments can also wield considerable influence over media coverage through financial and non-financial means (McMillan and Zoido, 2002; Besley and Prat, 2006; Tella and Franceschelli, 2011; Durante and Knight, 2012; Qian and Yanagizawa-Drott, 2017).

2013, 2017; Roberts, 2018) and leverage state ownership of media to ensure alignment with their political objectives (Djankov et al., 2003; Gehlbach and Sonin, 2014; Qin, Strömberg, and Wu, 2018; You, Zhang, and Zhang, 2018). We show that media bias against institutionally distant countries is significantly more pronounced in regimes with higher levels of censorship and state media ownership, and during periods of economic and political distress. These patterns underscore the role of international media coverage as a strategic tool of authoritarian information control, used not only to divert attention and shape domestic opinion, but also to export ideological narratives and erode the legitimacy of democratic rivals on the global stage.

Finally, our research contributes to a growing literature in economics that recognizes the importance of geopolitical forces for outcomes such as growth, trade, and capital flows (e.g., Caldara and Iacoviello, 2022; Clayton, Maggiori, and Schreger, 2023; Korovkin and Makarin, 2023; Korovkin, Makarin, and Miyauchi, 2025); for recent reviews, see Mohr and Trebesch (2024) and Clayton, Maggiori, and Schreger (2025). By documenting how differences in political institutions shape systematic variation in the tone and content of cross-national media coverage, our findings offer a new lens on how political and ideological alignments are reflected, and potentially reinforced, through global information flows.[4] Building on this insight, we use patterns in global media coverage to introduce a novel, granular measure of ideological and political alignment—contributing to recent efforts to capture bilateral alliances, rivalries, and spheres of influence (Liu and Yang, 2024; Broner et al., 2025).

The rest of the paper is organized as follows. Section 2 describes the data and key variables. Section 3 studies media bias in tone in coverage of institutionally distant countries. Section 4 analyzes bias in theme selection. Section 5 introduces a new metric of ideological alignment based on media tone co-movement. Section 6 assesses the economic and political relevance of global media bias. Section 7 concludes.

2 Data

2.1 Sample

Global Media Data. The data in this study are sourced from the Global Database of Events, Language, and Tone (GDELT) (Leetaru and Schrodt, 2013). The GDELT

[4]This perspective also complements work in international relations that emphasizes the role of soft power, strategic communication, and ideologically driven narratives in shaping state behavior and alliance structures (e.g., Jervis, 1976; Nye, 2004).

Project consists of two main databases: the GDELT Event Dataset and the GDELT Global Knowledge Graph (GKG). The Event Dataset, which focuses on recording *individual events* by aggregating data from various news articles, has been applied in recent economics research (e.g., Amarasinghe, 2022, 2023; Cantoni et al., 2024). Our research investigates cross-national media bias by analyzing how countries portray foreign events in domestic news coverage. As such, our analysis leverages the GKG Database 2.0, which records *individual news articles* and extracts granular metadata, allowing for a detailed analysis of their themes, contexts, and sentiments.

The database offers real-time translation of news articles from 65 languages and measurements of more than 2,300 emotions and themes. Using advanced computer algorithms and deep learning techniques, it extracts valuable information from billions of daily published words and images. Within 15 minutes of monitoring a breaking news report worldwide, the system translates the content and identifies the people, locations, organizations, themes, sources, emotions, quotes, and events. A detailed overview of the GKG Database 2.0 is provided in Online Appendix B, while two illustrative examples of articles from the database are provided in Online Appendix C.

Figure A1 provides an overview of the GKG Database 2.0, highlighting the global distribution of media outlets and the intensity of media coverage. Panel A illustrates the geographical distribution of media outlets included in our sample. Darker-blue areas represent countries with a higher concentration of media outlets, while lighter-blue areas indicate countries with fewer outlets. The United States stands out as the country with the highest number of media outlets, whereas many African countries have a limited media-outlet presence in the database. Panel B depicts the intensity of media coverage across the globe. Countries receiving more frequent news coverage are shaded in darker blue. Notable examples include the United States, the United Kingdom, and China, which receive extensive global media attention. In contrast, lighter-shaded areas represent countries that are covered less frequently.

We further illustrate the bilateral reporting relations of the database in Figure 1. Panel A presents a heat map visualizing the intensity of news coverage between countries. For simplicity, we focus on the top 50 countries ranked by the total number of news articles published by their media outlets. The y-axis represents the source country (the country of origin for the news article), and the x-axis represents the target country (the country being reported on). The color gradient indicates the volume of news articles originating from source countries and covering events or issues in target countries. Panel B shows a network figure where the size of each node corresponds to the total number of news articles published by that country's media outlets. The thickness of the edges reflects the volume of news coverage exchanged between countries. Green nodes represent liberal democratic countries, while red nodes represent non-liberal-democratic countries (including electoral democracies such as Poland). Together, Figures A1 and 1 demonstrate the database's extensive coverage, enhancing its credibility as a valuable resource for studying global media dynamics.

Countries of origin for media outlets. A challenging task we encountered was accurately identifying the source country of each media outlet. To address this challenge, we adopted a multistep approach. First, we leveraged a comprehensive Wikipedia list of 187 news agencies by country to classify major outlets. Second, we analyzed top-level domain suffixes (TLDs) to determine country affiliation (though this method was limited by the frequent use of generic domains: .com, .net, .org). For outlets not identified through these methods, we employed OpenAI's GPT-4o model to determine the country of origin. To evaluate its accuracy, we manually coded the country of origin for a random sample of 1,000 outlets and compared it to GPT-4o's classifications. The GPT-4o model achieved an impressive 94% accuracy rate in this validation exercise. A detailed discussion of our process is provided in Online Appendix B, while Online Appendix D lists the major media outlets identified. Finally, by combining the source country of each media outlet with the target country of its reporting, we merged the GDELT data with country-level political and economic indicators.

Themes. To classify the news articles into different themes, we rely on the World Bank Group (WBG) Topical Taxonomy, which has been integrated into the GDELT GKG 2.0. The WBG Topical Taxonomy comprises a broad spectrum of level-1 themes, including economic growth, poverty, trade, social development, education, health, environment, energy, public sector management, gender, and violence. Making the classification more granular, each level-1 theme is further divided into subthemes. In our research, we utilize 24 level-1 themes from the WBG Topical Taxonomy as reported in the GKG 2.0 database. Figure A2 presents the distribution of the themes. The themes most frequently covered in news articles are "Public Sector Management," "Health, Nutrition, and Population," and "Fragility, Conflict, and Violence." In contrast, themes such as "Climate Change," "Macroeconomic and Structural Policies," and "Public-Private Partnerships" appear far less frequently in the dataset. While certain themes feature a relatively higher volume of news articles, the distribution remains sufficiently dispersed to facilitate rigorous empirical analysis.

Sample construction. The sample we use in our main analysis was constructed through the following steps. First, we obtained raw news data from the GKG 2.0 database for the years 2015 to 2022, accessible via the GDELT website. This initial dataset comprised approximately 1.4 billion news articles (15.5 TB in total). After processing the metadata of all articles, we retained only those that reported on a single country and contained unique URLs. Next, we aggregated the data at the level of media outlet by target country by theme by month. At this level, we calculated the average tone of news coverage. To focus specifically on cross-border reporting, we excluded observations covering domestic news.

In our final sample, the tone in each observation represents the average sentiment of news coverage produced by a media outlet when reporting on a specific country during a given month and for a specific theme. The final sample comprises 81,036,750 observations, with 55,240 media outlets from 176 source countries reporting on 177 target countries between 2015 and 2022. The sample-construction process is detailed in Online Appendix B.

We construct our sample at the outlet-by-target-country-by-theme-by-month level for two primary reasons. First, as detailed below, we aggregate the sentiment measure of the news articles at this level to facilitate estimation. In contrast, the article-level dataset is prohibitively large, rendering econometric estimation computationally impractical. Second, this sample structure enhances identification by allowing econometric models to incorporate granular fixed effects relevant to our analysis. Specifically, by including target country $\times$ theme $\times$ year fixed effects, we examine within-theme variation in bias for a given target country in the same year. Furthermore, we account for media-outlet-level trends to further disentangle the relationship of interest.

2.2 Measuring Media Bias

Media bias can take on multiple forms, one of which is characterized by the "tone" and "sentiment" in news coverage (Puglisi and Snyder, 2015). In this paper, we focus primarily on the tone that one country's media outlets use to report on another country. We consider media coverage biased if it portrays another country in a *systematically* favorable or unfavorable manner.

As Puglisi and Snyder (2015) emphasize in their review, the term *bias* in the empirical literature on media rarely denotes a deviation from an objective ground truth, as such a benchmark is seldom observable and may even be impossible to define. Instead, it most often refers to bias relative to other outlets (Puglisi and Snyder, 2015, p. 648). To stay consistent with this usage, we employ the term *bias* throughout, though our analysis can equally be framed as an examination of *slant*.

Baseline Measure. The GKG 2.0 database provides three metrics for the baseline article tone: *Positive Score*, *Negative Score*, and *Tone*. For a given article q,

$$\text{Positive}_q = \left(\frac{\#\text{Pos}_q}{\#\text{Words}_q}\right) \times 100, \quad \text{Negative}_q = \left(\frac{\#\text{Neg}_q}{\#\text{Words}_q}\right) \times 100 \tag{1}$$

$$\text{Tone}_q = \text{Positive}_q - \text{Negative}_q \tag{2}$$

Here, $\#\text{Pos}_q$ and $\#\text{Neg}_q$ represent the number of positive and negative words, respectively, in article q, and $\#\text{Words}_q$ is the total number of words in article q. The *Tone* score ranges from -100 (extremely negative) to $+100$ (extremely positive), with most values between -10 and $+10$, and 0 indicating a neutral tone.

Validation. To validate the GKG 2.0 tone measures, we randomly sample 50,000 English-language articles per month from the GDELT GKG 2.0 database and use the article URLs (when available) to download the original content. We ensure that the number of words in the downloaded articles closely matches the word count reported by GDELT (with less than a 5% difference), yielding a dataset of 157,779 news articles. We then construct alternative tone measures using three standard dictionaries: the AFINN lexicon (Nielsen, 2011), the Bing Liu lexicon (Hu and Liu, 2004), and the NRC lexicon (Mohammad and Turney, 2013). We make adjustments to ensure the comparability of the alternative sentiment measures with the GDELT Tone by dividing all the metrics by the total word count of each article.

In Figure A3 of the Online Appendix, we compare the tone measures we derived with those provided in the GKG 2.0 database; we find significant Pearson correlations, ranging from 0.66 to 0.81.[5] These results suggest that the tone measures in the GKG 2.0 database are both informative and valid—they capture sentiment infor-

[5] As noted in the GKG 2.0 database, the GDELT team preprocesses news articles by filtering out content unrelated to the news itself. However, because their internal algorithm is not publicly available, we rely on the text directly downloaded from their website when constructing variables. This practice inevitably introduces some degree of noise.

mation comparable to algorithms widely used in the literature. Also, GKG 2.0 offers alternative tone measures based on the Lexicoder Sentiment Dictionary (LSD) and SentiWordNet 3.0 (SWN). We also demonstrate that our results remain highly robust to these alternative measures.

We aggregate the tone measure at the level of the media outlet, theme, and month to facilitate analysis using the baseline sample:

$$\text{Tone}_{i,j,s,m} = \frac{\sum_{q \in \Omega_{i,j,s,m}} \text{Tone}_q}{\text{N}_{i,j,s,m}} \tag{3}$$

where $\text{Tone}_{i,j,s,m}$ denotes the average tone for news articles q published by media outlet i, covering theme s, and referencing country j during year-month m, where $\text{N}_{i,j,s,m}$ reflects the number of articles in each $\{i, j, s, m\}$ cell.

Summary statistics and descriptive patterns. The summary statistics of the baseline sample are presented in Panel A of Table A1. The mean and the median of the tone measure are -1.929 and -1.677, respectively. The sentiment variable demonstrates a large variation with a standard deviation of 3.757. On average, the tone is negative, reflecting the higher proportion of negative words relative to positive words.

Figure A4 summarizes the distribution of tone. Panel A shows an approximately normal density. Panel B plots the mean and confidence intervals over time; despite substantial cross-sectional variation, these do not display a pronounced trend.

Figure A5 illustrates how average tone varies and evolves across prominent country pairs. Panel A shows that U.S. outlets portray Russia far more negatively than either the United Kingdom or China. Panel B reveals that Chinese outlets initially used similar tone toward the United States, United Kingdom, and Russia, but starting in 2018—the beginning of the U.S.-China trade war—their tone toward the United States became increasingly negative while coverage of Russia grew more favorable.[6]

Table A2 further reports summary statistics by selected country pairs and themes. Average tone displays intuitive patterns across geopolitical allies and adversaries (Panel A) and across thematic categories, with the most negative mean tone observed for the theme "Fragility, Conflict, and Violence" (Panel B).

Figure 2 focuses on the Russia-Ukraine dyad. Average media tone in Ukrainian outlets toward Russia dropped precipitously following the onset of the Russia-Ukraine

[6]The tone measures in Figure A5 are residualized by theme-by-year-month fixed effects to account for variation in topic composition over time. Patterns are similar without such residualization.

war in early 2022. In contrast, Russian media sentiment toward Ukraine declined more modestly over the same period. This asymmetry is consistent with well-documented features of Russian state-led media wartime discourse, including the avoidance of explicitly negative language, such as the use of the term "special military operation" and restrictions on the use of the word "war."[7] This illustrative example provides further validation that article tone constitutes a meaningful sentiment measure and that the selected sample is appropriate for our study.

2.3 Institutional Distance

To measure regime type, we use the Liberal Democracy Index from the V-Dem project (Coppedge et al., 2023). Each indicator in V-Dem is independently rated by at least five country experts (Pemstein et al., 2022). The index ranges from 0 to 1, with higher values indicating stronger adherence to liberal democratic principles such as the rule of law, checks and balances, and civil liberties.

Let $LDI_{A,t}$ and $LDI_{B,t}$ denote the liberal democracy scores of countries A and B in year t, respectively. We define Institutional Distance$_{AB,t} = |LDI_{A,t} - LDI_{B,t}|$ as our main independent variable. Panel B of Table A1 reports its summary statistics, and Figure A6 explores its distribution. As shown in Panel A of Figure A6, institutional distance varies widely across country pairs. Most country pairs have relatively small institutional distances, while their frequency declines as distance increases. Panel B indicates that, although institutional distance is not highly volatile, there is still meaningful time-specific variation once we residualize it by country-pair fixed effects. Further, a certain number of countries experience shifts in regime-type classifications in the V-Dem dataset during our study period. We use these cases, indicated in Table A17, when conducting an event-study analysis in Section 3.4.

In addition, we construct alternative measures of institutional distance by using the Electoral Democracy Index from the V-Dem project and the Polity Index from Marshall, Gurr, and Jaggers (2019). The Electoral Democracy Index primarily captures the core principles of electoral or representative democracy, focusing on the extent to which political leaders are elected through comprehensive voting rights in free and fair elections, as well as the degree to which freedoms of association and expression are safeguarded. In comparison, the Liberal Democracy Index takes into

[7] These dynamics also echo asymmetric public opinion responses observed during earlier stages of the conflict (Korovkin and Makarin, 2023), as well as sharp shifts in the language used in Ukrainian social media following the full-scale invasion (Abramenko, Korovkin, and Makarin, 2024).

account broader dimensions of democracy, such as civil liberties and checks and balances, providing a more comprehensive assessment of political institutions. Though the Polity Index is widely used, it is updated only through 2018. As part of our robustness checks, we confirm that our results remain consistent when using these two alternative measures of democracy.

2.4 Other Data

We retrieve datasets from several sources to construct variables that fall into four categories: (a) country-pair and year level; (b) country-year level (c) firm-specific variables, and (d) public opinion across countries.

Country-pair and year level variables. As an alternative measure of political and ideological alignment, we use ideal point distances in UN General Assembly (UNGA) voting, which capture the extent to which countries share similar foreign policy preferences (Voeten, 2000, 2013; Bailey, Strezhnev, and Voeten, 2017). These data are obtained from United Nations General Assembly Voting Data on Harvard Dataverse (Voeten, Strezhnev, and Bailey, 2023). To control for linguistic similarity, we use the Common Official Language (COL) indicator (Melitz and Toubal, 2014), sourced from the Centre d'Études Prospectives et d'Informations Internationales (CEPII). We also use a measure of religious distance developed by Dow and Karunaratna (2006). This composite measure combines three 5-point-scale items into a single factor. The first item is the distance between the two closest major religions for each pair of countries. The second and third items measure the incidence of one country's major religion(s) in other countries. These three items are then collapsed into a unified factor score. We calculate the geographic distance between countries using the geodesic distance between their capital cities, relying on a dataset compiled by researchers affiliated with CEPII. Finally, genetic distance between countries is measured by the distance to "the most recent common ancestors of two populations, or, equivalently, of their degree of genealogical relatedness" (Spolaore and Wacziarg, 2009, p.481). The summary statistics of these variables are presented in Panel B of Table A1.

Country-year-level variables. To investigate the mechanisms, we use the source country's monthly GDP growth rate measured with night-light data from the the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) dataset and inflation data taken from the World Development Indicators (WDI) database. To measure the extent of government control over the media in the source country, our analysis uses indicators of media censorship and state media ownership derived from the V-Dem database (Coppedge et al., 2023). Media censorship indicators include "government censorship effort" and "internet censorship effort," whereas media state ownership indicators include "state-owned print media" and "state-owned broadcast media." The two media-censorship indicators assign higher values to indicate lower levels of censorship in the V-Dem database. To avoid confusion, we rename these indicators as "traditional media freedom" and "internet freedom," respectively. The summary statistics of these variables are presented in Panel C of Table A1.

Public opinion across countries. We also explore the potential political relevance of global media coverage by examining whether changes in media tone toward a foreign country are associated with subsequent shifts in public attitudes toward that country. Our analysis draws on the Gallup World Poll (GWP), a global survey conducted annually since 2005 across more than 160 countries covering 99% of the world's population. The GWP tracks a wide range of issues—including employment, governance, and well-being—through over 100 standardized questions, enabling consistent trend analysis and cross-country comparisons. The GWP dataset has been used to study political approval and confidence in government across countries (e.g., Guriev and Treisman, 2020a; Guriev, Melnikov, and Zhuravskaya, 2021). Following Goldsmith, Horiuchi, and Matush (2021), we draw on GWP data to evaluate public favorability toward foreign countries, using the question "Do you approve or disapprove of the job performance of the leadership of (the name of a country)?" Since respondents often have limited knowledge of foreign countries' actual governance, their answers likely reflect the overall favorability or the broader "image" of the country and its political regime.

3 Institutional Distance and News Coverage Tone

3.1 Baseline Specification

We seek to establish the presence of global media bias by examining whether the tone of a country's media toward foreign nations varies systematically with institutional distance. Following the logic of gravity models in international trade (Anderson and Van Wincoop, 2003), we estimate the following baseline specification:

$$\text{Tone}_{i(c),j,s,m(t)} = \beta\,\text{Institutional Distance}_{c,j,t-1} + \theta_{j,s,t} + \delta_{i(c),s,t} + \epsilon_{i(c),j,s,m(t)} \tag{4}$$

where $\text{Tone}_{i(c),j,s,m(t)}$ denotes the average tone of articles published by media outlet i (based in source country c) about target country j, within thematic category s and year-month m (of year t). The key independent variable is Institutional $\text{Distance}_{c,j,t-1}$, defined as the absolute difference in liberal democracy scores between countries c and j in year $t-1$, drawn from V-Dem (for details, see Section 2.3).[8]

The richness of the dataset allows us to include two high-dimensional sets of fixed effects, mirroring multilateral resistance terms in the gravity literature (Larch and Yotov, 2024). $\theta_{j,s,t}$ absorb all time-varying variation specific to the target country, theme, and year, such as political shocks or major events that affect how a given country is covered on a particular issue. $\delta_{i,s,t}$ are the outlet-by-theme-by-year fixed effects, capturing variation in overall editorial stance or coverage intensity across themes and years at the outlet level. As such, the identifying variation comes from year-to-year changes in *relative* institutional distance between countries—net of any uniform shifts in tone associated with the *absolute* democracy levels of the source or target country within a given theme-year. To adjust for serial correlation, standard errors are clustered at the country-pair level.

For the β estimate to have a causal interpretation, the key identifying assumption is that, conditional on outlet-by-theme-by-year and target-country-by-theme-by-year fixed effects, changes in institutional distance between the source and target countries are orthogonal to unobserved factors that also influence media tone. That is, one would need to assume that the variation in institutional distance captures shifts in ideological and geopolitical divergence rather than being systematically driven by contemporaneous shocks that independently affect tone. To the extent that these contemporaneous shocks capture broader geopolitical frictions—such as diplomatic disputes or bilateral crises—this remains consistent with our conceptual framework, which interprets institutional distance as also a proxy for geopolitical and ideological misalignment. We further demonstrate the robustness of our baseline results to controls for trade flows, even-more-demanding sets of fixed effects, and other sensitivity checks that together rule out many of the potential omitted-variable concerns.

That said, we remain cautious about making strong causal claims from the baseline specification. Our primary goal is to test whether institutional distance and media tone are systematically linked even under demanding sets of controls. Establishing

[8] Using a lagged measure ensures that our estimates are not mechanically affected by contemporaneous media coverage, but results are similar when using contemporaneous institutional distance.

such a robust conditional relationship is already highly informative on its own and, in our view, sufficient to document systematic global media bias along ideological and geopolitical lines. However, to provide further evidence consistent with a causal interpretation, we complement the baseline analysis with a DID design that exploits episodes of democratic erosion in the source country (see Section 3.4).

3.2 Baseline Results

Table 1 presents the baseline results. Across columns (1)–(3), we introduce increasingly stringent fixed effects. Column (1) includes target-country × theme × year and outlet fixed effects; column (2) adds outlet × year fixed effects; and column (3) implements our preferred specification from Equation (4), incorporating both outlet × theme × year and target-country × theme × year fixed effects.

The coefficient on institutional distance remains stable and statistically significant across specifications. The estimate presented in column (3) indicates that an increase in institutional distance from zero to one is associated with a 0.260 decrease in the average media tone between two countries, which corresponds to 13.5% of the dependent variable mean and only slightly smaller than the drop in tone observed in Russia's coverage of Ukraine following the onset of the full-scale war in February 2022, in Figure 2. The magnitude of this media bias is amplified under certain conditions, as demonstrated by the subsequent heterogeneity tests.

Overall, these results point to a robust negative association between institutional distance and international media sentiment: countries with more-dissimilar political institutions tend to systematically report on one another in more negative terms. This pattern is consistent with ideologically driven bias in international news coverage.

Negative sentiments and emotions. Negative sentiment plays a central role in propaganda and is likely to be salient in the portrayal of hostile countries. Negative tone and emotions directed at a common "enemy" can strengthen in-group solidarity and national identity (e.g., Weiss, 2014). By defining an out-group, propaganda reinforces in-group boundaries and renders internal conflicts comparatively less salient.

To examine the role of negative sentiment in global media coverage, we disaggregate the tone measure into positive and negative components. Consistent with this hypothesis, the results in Table A3 show that negative sentiment is more strongly associated with institutional distance than positive sentiment, even after accounting for differences in their respective distributions.

We further investigate the role of negative emotions using fine-grained sentiment measures from GDELT based on the WordNet-Affect (WNA) lexicon (Strapparava and Valitutti, 2004), which provides detailed affective annotations of emotional content in text. For each article, we compute WNA scores as the share of words linked to each emotional category, then we aggregate these measures at the outlet–target-country–theme–month level. Figure A7 reports estimates from our baseline specification (4), now linking institutional distance to a range of WNA sentiment categories. Greater institutional distance is associated with higher frequencies of negative emotions, such as fear, sadness, and anxiety, and with lower frequencies of positive emotions, such as liking and love. These findings point to a systematic negative emotional coloring of media coverage between ideologically and politically divergent countries.

Baseline results by year and by theme. Online Appendix E reports the baseline estimates by year and theme. Consistent with an increasingly contentious geopolitical environment over the study period, the relationship between media tone and institutional distance becomes more negative over time (Table A5). Media bias is also consistently present across broad thematic categories, with larger coefficients for themes directly related to government activity (Figure A8).

Alternative alignment measure. One may argue that institutional distance captures only a subset of the relevant variation in political alignment over time, as democracies and autocracies can at times find themselves aligned geopolitically. In Panel A of Table A4, we examine media bias using an alternative measure of geopolitical alignment: ideal point distance in UNGA voting decisions (Voeten, 2000, 2013; Bailey, Strezhnev, and Voeten, 2017). Consistent with the notion that geopolitical divergence shapes international news coverage—and echoing our baseline results—we find that greater voting distance is associated with a more negative average media tone. In Panel B, we jointly include both alignment measures simultaneously and further control for time-invariant indicators of bilateral similarity, such as linguistic, religious, genetic, and geographic distances. In this horse-race specification, the coefficient on institutional distance remains statistically significant and only slightly attenuated, while the coefficient on UN voting distance becomes insignificant and shrinks considerably in magnitude.[9]

Taken together, these results suggest that institutional distance captures distinct

[9]The coefficients on institutional and UN voting distances remain similar when excluding the additional bilateral controls.

and substantively meaningful dimensions of ideological and political divergence that are particularly salient for international media tone, while UNGA voting distance appears less informative in this context despite its widespread use in the international relations and geoeconomics literature.

Robustness. We conduct an extensive battery of robustness checks; full details are available in Online Appendix E.

Our baseline findings are robust to a wide range of alternative specifications and measures. The negative relationship between institutional distance and media tone remains significant when including country-pair fixed effects and when adopting even-more-demanding monthly fixed-effects structures. The results are also robust to using alternative sentiment measures and alternative measures of political institutions.

Addressing concerns related to machine translation, we show that tone is not systematically correlated with translation status, that results are stable when controlling for the share of English-language articles, when restricting the sample to countries with widely spoken official languages, and when excluding major languages one at a time. The estimates are likewise stable across alternative rules for assigning outlets to source countries, including domain-based and manual classifications; if anything, focusing on large, well-known news agencies yields an even stronger relationship.

The results are further robust to alternative samples and inference choices. Estimates remain similar when winsorizing tone, excluding the United States, aggregating the data to higher levels, or sequentially dropping influential country pairs or random subsamples. Replicating the analysis using the GDELT Event 2.0 Dataset instead of GKG 2.0 yields comparable estimates, despite differences in classification structure. Controlling for bilateral economic interdependence and conflict—such as trade exposure, trade balances, and sanctions—does not materially affect the coefficient on institutional distance. Finally, statistical significance of our estimates is unaffected by alternative clustering schemes, including two-way clustering at different levels.

3.3 Biased News Coverage on National Leaders

Although the breadth of GKG 2.0 article themes is a strength—allowing us to map out the entire available global media landscape and show that ideologically driven global media bias extends beyond overtly political topics—one may expect such bias to be particularly strong in explicitly political coverage. We therefore focus on news articles about national leaders, who personify their countries on the global stage and are closely associated with their regimes' political actions (Saich, 2014). Using a curated sample of leader-related articles (for details, see Online Appendix B.2), we estimate a series of increasingly demanding specifications that compare how different outlets cover the same leaders while flexibly controlling for outlet- and leader-specific time variation.[10] The results are reported in Table 2.

Across all specifications in Table 2, institutional distance remains strongly and statistically significantly associated with a more negative media tone, with effect sizes more than three times larger than in the baseline results reported in Table 1. In our preferred specification, mirroring the fixed effects structure of the baseline equation, increasing institutional distance from zero to one is associated with a 0.916-point decline in average tone. This corresponds to approximately 26.8% of the standard deviation and roughly *two-thirds* of the drop observed in average Ukrainian outlets' coverage of Russia following the 2022 invasion. These results indicate that global media bias is particularly pronounced in highly salient political contexts, such as the coverage of national leaders in more institutionally distant countries.

3.4 Democratic Erosion as a Natural Experiment

We exploit episodes of democratic erosion in the source country to provide further evidence that institutional distance plays a role in shaping global media coverage.

Using the Regimes of the World (RoW) classification from V-Dem (Lührmann, Tannenberg, and Lindberg, 2018), we define democratic erosion as a downward shift in regime type—from liberal democracy to electoral democracy, from electoral democracy to electoral autocracy, or from electoral autocracy to closed autocracy. To account for early-stage backsliding prior to formal reclassification, we draw on the Episodes of Regime Transformation (ERT) dataset (Maerz et al., 2023) and exclude cases in which erosion began before 2015, the start of our sample. We also exclude countries that experienced democratic transitions *during* the period; our exercise yielded 17 democratic erosion episodes (listed in Table A17).

If institutional distance—and the political and ideological divergence it captures—

[10] Given the smaller number of articles used in this analysis relative to the baseline sample, it is now feasible to estimate a version of Equation (4) at the article level. Specifically, we estimate:

$$\text{Tone}_{q(i,c),l(j),m(t)} = \beta\, \text{Institutional Distance}_{c,j,t-1} + \theta_{l(j),t} + \delta_{i(c),t} + \epsilon_{q(i,c),l(j),m(t)} \quad (5)$$

where $\text{Tone}_{q(i,c),l(j),m(t)}$ denotes the tone of article q, published by outlet i in country c, covering national leader l of country j in month m of year t; and $\theta_{l(j),t}$ and $\delta_{i(c),t}$ represent leader–year and outlet–year fixed effects, respectively. Standard errors are clustered at the country-pair level.

drives media bias, then democratic backsliding should widen a country's distance from democracies and lead its media to portray democratic countries more negatively. We test this implication by estimating the following staggered DID equation:

$$
\begin{aligned}
\text{Tone}_{i(c),j,s,m(t)} = &\sum_{\substack{\tau=-4 \\ \tau\neq-1}}^{4} \beta_\tau \left(\text{Erosion}_{c,\tau} \times \text{Liberal Democracy Indicator}_{j,t}^{\text{Target Country}}\right) \\
&+ \delta_{i(c),s,t} + \theta_{j,s,t} + \alpha_{i(c),j} + \epsilon_{i(c),j,s,m(t)}
\end{aligned} \tag{6}
$$

where $\text{Erosion}_{c,\tau}$ is a set of indicator variables that equal one if τ periods have passed since democratic erosion in source country c, with period $\tau = -1$ serving as the baseline, and Liberal Democracy $\text{Indicator}_{j,t}^{\text{Target Country}}$ is a binary indicator that equals one if the target country j is classified as a liberal democracy in year t (or, alternatively, a continuous Liberal Democracy Index). We estimate dynamic effects using the Sun and Abraham (2021) estimator and report OLS estimates for comparison.

Figure 3 plots the event-study coefficients, with full results reported in Table A19. Media tone toward democratic countries becomes significantly more negative beginning roughly two years after democratic erosion, with no evidence of pretrends. A nondynamic specification (Table A18) implies a 0.089-point decline in tone toward democratic countries, corresponding to 2.4% of the outcome's standard deviation. The results are robust to excluding individual erosion episodes (Figure A11), ensuring that the estimates are not driven by any particular tangential case. Overall, these findings further support the role of institutional distance in shaping global media bias.

3.5 Heterogeneity and Mechanisms

The strength of global media bias varies systematically across political and informational environments, shedding light on the mechanisms behind our baseline results.

3.5.1 Democracy in Source and Target Countries

We first show that media bias is *driven primarily by less-democratic source countries* and is disproportionately *directed at democratic targets*. We augment our baseline specification (Equation (4)) with an interaction between institutional distance and an indicator for whether the source country is classified as a liberal democracy. We then use an analog model to study the heterogeneity of our baseline estimates by whether the target country is the liberal democracy.

In columns (1) and (2) of Panel A in Table 3, we examine how the degree of political-regime-driven media bias varies depending on whether the source country is a liberal democracy. Column (2) adopts the more detailed baseline specification, with media-outlet × theme × year and target-country × theme × year fixed effects. In both cases, the results indicate that the negative media bias toward more institutionally distant countries is largely mitigated when the source country is a liberal democracy. Conversely, the relationship is more pronounced for outlets in less democratic countries. To put the magnitudes in perspective, the standalone coefficient on institutional distance in column (2) is –0.387, the interaction with the "liberal-democratic source" indicator is +0.246, implying a net estimate of –0.141 for liberal-democratic source countries—about 46% smaller than the baseline (–0.260 in Table 1).

Figure A12 illustrates these patterns. After residualizing tone by outlet × theme × year and target × theme × year fixed effects, Panels A and B plot the residuals against institutional distance separately for liberal democracies and less-democratic source countries. The slope is negative—and steeper beyond a distance of 0.5—for less-democratic source countries, but is essentially flat for liberal democracies.

Columns (3) and (4) of Panel A in Table 3 shift the focus to the target country's regime. When the target is a liberal democracy, the distance-tone relationship roughly doubles in magnitude relative to the baseline (–0.424 in column (4) versus –0.260 in Table 1); for less democratic targets it is smaller and only marginally significant (–0.106). Panels C and D of Figure A12 confirm this: the downward slope is noticeably steeper when the target is a liberal democracy, particularly once regime distance exceeds 0.5. In summary, media bias appears to be stronger in cases when the source country is less democratic but also, and even more strongly, for when the country being covered is a liberal democracy.

3.5.2 Information Control

Next, we show that the relationship between institutional distance and media tone is stronger in countries with *tighter control over information flows.*

To sustain regime stability, governments—particularly authoritarian ones—often employ information-control strategies to manipulate the flow of information (Guriev and Treisman, 2019, 2020b). These strategies include censorship, which suppresses dissenting voices and shapes the narrative and tone of media coverage and online discourse (King, Pan, and Roberts, 2013, 2017; Roberts, 2018). Authoritarian governments also actively promote state-owned media to ensure its dominance across critical information channels, including print media, television stations, and social media platforms (Gehlbach and Sonin, 2014). As a result, the media landscape in authoritarian countries tends to be characterized by higher levels of state ownership (Djankov et al., 2003). Government-controlled media outlets primarily serve political imperatives rather than commercial interests (Qin, Strömberg, and Wu, 2018; You, Zhang, and Zhang, 2018).

We examine how the extent of global media bias varies across countries with different levels of information control by augmenting our baseline specification with an interaction between institutional distance and various indicators detailed below.

Media freedom. Panel B of Table 3 explores whether the association between institutional distance and tone differs with the media freedom level in the source country. Columns (1) and (2) interact institutional distance with traditional-media freedom (TV, radio, print), while columns (3) and (4) use an internet-freedom measure.

Across all columns, the main coefficient on institutional distance—interpreted here as the estimated effect when media freedom is at its minimum—is negative and statistically significant. The interaction terms are positive and significant, indicating that the negative association between distance and tone is weaker where media freedom is higher. In other words, stricter controls on either traditional or online media are associated with a stronger negative bias against institutionally distant countries, whereas freer media environments show much smaller bias levels.

State media ownership. Panel C of Table 3 turns to media ownership. Columns (1) and (2) interact institutional distance with a measure of state ownership in print media; columns (3) and (4) do the same for broadcast media. In each case, the interaction term is negative and statistically significant, implying that outlets operating in countries with greater state media ownership tend to cover institutionally distant nations in a more negative light.

Table A20 offers a complementary test using a binary indicator for whether an outlet itself is state-owned (classified via GPT-4o). The baseline coefficient for non-state outlets is –0.23, while the interaction is about –0.29, suggesting a noticeably larger association for state-owned outlets.

Taken together, these patterns suggest that state media ownership in the source country is associated with relatively more-negative coverage of institutionally distant countries, consistent with stronger ideological and political media bias.

3.5.3 Distress

We next show that the negative association between institutional distance and media tone becomes stronger when *the source country faces economic headwinds*.

Prior work suggests that governments—especially autocratic ones—tighten informational controls in lean times. For instance, Martinez (2022) shows that autocracies overstate GDP growth when domestic conditions deteriorate, while Rozenas and Stukal (2019) document how Russian state television shifts blame for bad economic news onto foreign actors during downturns.

Because headline GDP figures can be misreported or manipulated, particularly in less-democratic or lower-capacity states, we proxy for economic performance with satellite-based night-light intensity (Henderson, Storeygard, and Weil, 2012). Using the VIIRS DNB data aggregated to the country-month level, we explore two sources of heterogeneity in Table 4.

Columns (1) and (2) interact institutional distance with monthly growth of the source country's night-light intensity. The coefficient on the interaction term is positive, suggesting that stronger domestic growth is associated with a weaker distance-tone relationship, whereas slower growth coincides with a more intensive media bias toward more institutionally distant countries. Columns (3) and (4) replace growth with an indicator for high inflation ($> 5\%$). Here the interaction is negative and significant, implying that elevated inflation is linked to a stronger negative relationship between institutional distance and tone.

3.5.4 Domestic Turmoil

Finally, we further show that biased coverage intensifies during periods of *domestic political instability* in less democratic source countries.

Following Amarasinghe (2022), we construct an index of domestic turmoil using all domestic events targeting the government and the associated sentiment on the Goldstein scale (Goldstein, 1992) from the GDELT Event 2.0 dataset. This index measures people's resentment toward their government by calculating the proportion of events with negative sentiment scores among all events targeted at the government.

Column (1) of Table 5 shows that, in the full sample, higher domestic turmoil correlates with a more negative overall media tone toward foreign countries. Column (2) adds an interaction term between institutional distance and domestic turmoil; the coefficient is imprecisely estimated, suggesting no clear differential pattern for the sample as a whole. However, focusing on source countries that are not liberal democracies in column (3), the interaction term is negative and statistically significant: when turmoil rises at home, outlets in less democratic countries tend to adopt a sharper negative tone toward institutionally distant countries—i.e., democracies. By contrast, column (4) finds no comparable association for outlets based in liberal democracies. Overall, these patterns are consistent with the idea that authoritarian regimes may rely more heavily on negative international coverage, particularly of democracies, when facing domestic unrest.

Taken together, the heterogeneity patterns point to a clear logic: media bias is most pronounced when the source country is less democratic and the target country is a liberal democracy, and it strengthens under conditions associated with tighter information control or heightened domestic pressure. These findings are consistent with changes in state-aligned media narratives in authoritarian settings that emphasize negative portrayals of democratic countries, particularly during periods when managing domestic perceptions may be especially important.

Illustrative examples are consistent with this interpretation. For instance, during the COVID-19 pandemic, Chinese state media emphasized perceived shortcomings in the U.S. pandemic response while highlighting China's capacity for rapid social mobilization (Xia, Huang, and Zhang, 2022; Fu, 2023).[11] Similarly, Syrian state media framed the Arab Spring uprisings as the result of foreign interference rather than domestic grievances (Alrababa'h and Blaydes, 2021). Such patterns align with prior evidence that externally focused narratives can foster nationalist sentiment and rally domestic support during periods of political strain (Weiss, 2014).

[11] See also: *Global Times,* March 6, 2020. "China outperforms the United States in public health," https://oversea.huanqiu.com/article/9CaKrnKpKEV?bsh_bid=5488851275 (accessed May 21, 2025); and *Global Times,* April 9, 2020. "The US has dragged down the fight against the epidemic, and it has also dimmed the prospects of the world," https://opinion.huanqiu.com/article/3xlaqBQqzYD (accessed May 21, 2025).

4 Institutional Distance and News Theme Selection

Media bias may operate not only through systematic differences in tone but also through selective topic choice, whereby outlets emphasize themes that typically carry negative connotations. For instance, Chinese media may disproportionately highlight incidents of gun violence when reporting on the United States, and American media may spend a disproportionate amount of time covering the degree of state surveillance in China compared to other countries. To test this mechanism, we estimate the following specification:

$$\begin{aligned} \text{Theme Share}_{c,j,s,t} = \beta &\left(\text{Institutional Distance}_{c,j,t-1} \times \overline{\text{Theme Tone}}_s\right) \\ &+ \gamma \, \text{Institutional Distance}_{c,j,t-1} + \delta_{j,s} + \theta_c + \alpha_t + \epsilon_{c,j,s,t} \end{aligned} \tag{7}$$

where $\text{Theme Share}_{c,j,s,t}$ is the share of news in theme s scaled by total number of news articles by source country c covering target country j in year t. In alternative specifications, the dependent variable is a binary indicator for whether theme s is covered at all. Given the computational constraints, we aggregate the data at the source-country–target-country–theme–year level.

$\overline{\text{Theme Tone}}_s$ measures the average sentiment associated with theme s and is constructed in two ways. First, we compute the average tone across all articles classified under theme s. Second, we compute a target-specific measure by averaging the tone of all articles covering target country j, theme s, and year t, excluding articles published by outlets from source country c. We expect the interaction coefficient β to be negative: when a theme is inherently negative, greater institutional distance should be associated with a higher share of coverage devoted to that theme.

Columns (1) and (2) of Table 6, which rely on the two alternative measures of theme tone, confirm this prediction. Columns (3) and (4), which use a binary indicator for whether a theme is covered, yield consistent results. Column (5) focuses on the "violence" theme—one of the most negative topics in our sample—and shows that greater institutional distance significantly increases the share of violence-related coverage. Taken together, these findings indicate that as institutional distance grows, outlets systematically emphasize more-negative themes when reporting on foreign countries. Thus, ideologically driven global media bias manifests not only in the tone of coverage but also in topic selection.

5 Media Alliances in Ideological Warfare

Our findings thus far show that media outlets portray institutionally similar countries more favorably. We now broaden the analysis by using media narratives to uncover *media alliances* in a broader ideological contest among geopolitical blocs. To do so, we construct a correlation-based alignment metric that captures how similarly two countries' media outlets report on *third* countries across themes. This measure is granular, high-frequency, and scalable, and provides a novel proxy for ideological proximity, responding to recent calls for finer measures of "geoeconomic" alignment (Mohr and Trebesch, 2024).

We begin by aggregating the baseline tone measure to the source-country–target-country–theme–year level:

$$\text{Tone}_{\substack{c,j,s,t \\ j\neq c}} = \frac{1}{\text{N}_{c,j,s,t}} \sum_{(i,m)\in\Omega_{c,j,s,t}} \text{Tone}_{i(c),j,s,m(t)} \tag{8}$$

where $\text{Tone}_{i(c),j,s,m(t)}$ is the average tone of news from media outlet i in country c when reporting country j for theme s during month m of year t and $\text{N}_{c,j,s,t} = |\Omega_{c,j,s,t}|$ is the count of news within each $\{c, j, s, t\}$ cell. To ensure the stability of the tone measure, we require that each such country-pair–theme–year cell contains at least 12 observations, i.e., at least one article per month on average.

For each pair of countries (A, B) in year t, we then compute the Pearson correlation between their tone measures toward the same themes in the same *third* countries, excluding coverage of each other. This yields an annual country-pair alignment measure, $\rho_{A,B,t}$, calculated using samples with at least 100 observations to ensure reliability.

Panel A of Figure 4 plots the residualized correlation coefficient, adjusted for source-country, target-country, and year fixed effects, against institutional distance. The clear negative slope indicates that countries with similar political regimes tend to cover third-country events in a synchronized manner, suggesting the presence of narrative or media alliances between countries.

Panel B of Figure 4 visualizes the alliance network: each country is a node (with size proportional to its news volume), and an edge links two countries if their average raw correlation value between 2015 and 2022 falls in the top quintile. With information of regime types in 2015, blue nodes denote liberal democracies while red nodes denote less democratic countries. Liberal democracies form a tight cluster, with

Great Britain and especially the United States broadly at the center of the global media landscape. By contrast, less democratic countries form separate, smaller hubs, with Russia and China located on the periphery of the main network. This structure mirrors the findings in Panel A: in ideological contests, countries appear to gravitate toward others with similar regimes and coordinate their media narratives accordingly.

Figure A13 traces the evolution of these networks over time. Panel A shows that connectivity between liberal democracies and less democratic countries—measured by the average edge weight across groups—declines sharply after 2019, coinciding with rising geopolitical fragmentation following the COVID-19 pandemic and the Russia-Ukraine war. Panel B plots the clustering coefficient within each group, revealing relatively stable internal cohesion among liberal democracies but a steady post-2017 decline among the less democratic countries. Panels C and D report degree centrality for key countries: the United States and the United Kingdom maintain consistently high centrality within the liberal bloc, while China (from 2019) and Russia (from 2020) experience marked declines in centrality within the nonliberal bloc. Together, these patterns illustrate the ability of the media-based alignment measure to capture both cross-sectional structure and dynamic changes in geopolitical narratives.

Finally, we assess whether and how our novel media-based measure of geopolitical and ideological alignment relates to other forms of international alignment. As shown in Table A21, higher media alignment in year $t-1$ significantly predicts UN voting similarity, the formation of bilateral investment treaties, and participation in regional trade agreements in year t. These results suggest that similarities in media narratives capture latent ideological and strategic affinities that precede—and may even help shape—formal diplomatic and economic ties.

6 Political Relevance of Media Bias

In our final analysis, we examine the potential political consequences of global media bias by studying its predictive relationship with public opinion toward foreign countries. Specifically, using Gallup World Poll data, we test whether the average media tone in source country i toward target country j in year $t-1$ predicts public opinion in country i toward country j in year t. Public opinion is measured by via respondents' approval of the target country's leadership, based on the question "Do you approve or disapprove of the job performance of the leadership of (the name of a country)?"[12] Because respondents often have limited knowledge of foreign countries' actual governance, leadership approval serves as a meaningful proxy for broader sentiment toward the target country (Goldsmith, Horiuchi, and Matush, 2021).

Table 7 reports the results. Across specifications with increasingly demanding fixed effects, average media tone is positively associated with subsequent public approval of the target country's leadership. In the fully saturated specification, a one-standard-deviation increase in media tone toward a foreign country is associated with roughly a one-percentage-point increase in approval of that country's regime in the following year. These findings suggest that domestic media narratives may play an important role in shaping public perceptions of foreign countries.

7 Conclusion

This paper provides the first systematic evidence of global media bias: the tendency of national media outlets to portray foreign countries more negatively when they are ideologically and politically distant. Drawing on a corpus of 1.4 billion articles from over 55,000 outlets across 176 countries published between 2015 and 2022, we document that greater institutional distance—measured by differences in the countries' democracy scores—is associated with systematically more negative reporting, shaping both the sentiment and thematic content of news articles. This relationship persists after conditioning on a rich set of fixed effects and controls, as well as across a wide range of robustness checks.

We show that global media bias is particularly strong when the target is a liberal democracy and when the source is not. Bias also intensifies with higher levels of censorship and state media ownership, as well as during periods of economic or political distress. A difference-in-differences design exploiting episodes of democratic erosion supports a causal interpretation: as countries become less democratic, their media grow disproportionately more negative toward liberal democracies.

---

[12]Specifically, we estimate the following equation:

$$\mathrm{Y}_{i(c),j,t} = \beta\, \overline{\mathrm{Tone}}_{c,j,t-1} + \mathbf{X}_{i(c),t}\gamma + \delta_{c,t} + \theta_{j,t} + \epsilon_{i(c),j,t}$$

where $\mathrm{Y}_{i(c),j,t}$ is an indicator for whether respondent i from country c approves of the leadership of target country j in year t; $\overline{\mathrm{Tone}}_{c,j,t-1}$ is the average tone of news articles from country c about country j in year $t-1$; $\mathbf{X}_{i(c),t}$ is a vector of individual-level demographic controls. Standard errors are clustered at the country-pair level.

We also document the political relevance of cross-country news coverage: media tone predicts subsequent domestic public opinion toward foreign governments. This finding suggests that cross-country media coverage may influence mass perceptions and highlights its possible role as an underappreciated vector of soft power.

Beyond documenting bias, we introduce a new measure of international ideological alignment based on the co-movement of media tone across third countries. This method reveals an ideological clustering of countries: liberal democracies are tightly aligned in their media narratives, especially around the United States, while authoritarian regimes, such as China and Russia, form smaller alternative media blocs. This approach offers a scalable, high-frequency complement to more-slow-moving measures of international alignment, such as treaty networks or UN voting records.

Taken together, our results speak to the increasingly central role of information flows in the global economy. While trade, finance, and migration have long been studied through the lens of gravity models, we show that international narratives—how countries cover one another—also follow systematic patterns shaped by institutional proximity and political incentives. In doing so, our work opens a new empirical and conceptual frontier at the intersection of international economics, international relations, and political economy, with media-driven information flows as a key channel.

References

Abramenko, Serhii, Vasily Korovkin, and Alexey Makarin. 2024. "Social Media as an Identity Barometer: Evidence From the Russia-Ukraine War." In *AEA Papers and Proceedings*, vol. 114. 70–74.

Acemoglu, Daron, Tarek A. Hassan, and Ahmed Tahoun. 2018. "The Power of the Street: Evidence from Egypt's Arab Spring." *The Review of Financial Studies* 31 (1):1–42.

Alrababa'h, Ala' and Lisa Blaydes. 2021. "Authoritarian Media and Diversionary Threats: Lessons from 30 Years of Syrian State Discourse." *Political Science Research and Methods* 9 (4):693–708.

Amarasinghe, Ashani. 2022. "Diverting Domestic Turmoil." *Journal of Public Economics* 208:104608.

———. 2023. "Public Sentiment in Times of Terror." *Journal of Development Economics* 162:103058.

Anderson, James E and Eric Van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review* 93 (1):170–192.

Ash, Elliott, Ruben Durante, Maria Grebenschikova, and Carlo Schwarz. 2021. "Visual Representation and Stereotypes in News Media." .

Baccianella, Stefano, Andrea Esuli, and Fabrizio Sebastiani. 2010. "Sentiwordnet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining." In *Lrec*, vol. 10. 2200–2204.

Bachman, Jeffrey S and Esther Brito Ruiz. 2024. "The Geopolitics of Human Suffering: A Comparative Study of Media Coverage of the Conflicts in Yemen and Ukraine." *Third World Quarterly* 45 (1):24–42.

Bailey, Michael A., Anton Strezhnev, and Erik Voeten. 2017. "Estimating Dynamic State Preferences from United Nations Voting Data." *Journal of Conflict Resolution* 61 (2):430–456.

Baum, Matthew A and Yuri M Zhukov. 2015. "Filtering Revolution: Reporting Bias in International Newspaper Coverage of the Libyan Civil War." *Journal of Peace Research* 52 (3):384–400.

———. 2019. "Media Ownership and News Coverage of International Conflict." *Political Communication* 36 (1):36–63.

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

Broner, Fernando, Alberto Martin, Josefin Meyer, Christoph Trebesch, and Jiaxian Zhou Wu. 2025. "Hegemony and International Alignment." .

Caldara, Dario and Matteo Iacoviello. 2022. "Measuring Geopolitical Risk." *American Economic Review* 112 (4):1194–1225.

Cantoni, Davide, Andrew Kao, David Y. Yang, and Noam Yuchtman. 2024. "Protests." *Annual Review of Economics* 16.

Caprini, Giulia. 2023. "Visual Bias." *Working Paper* .

Chen, Heng and Li Han. 2022. "Do the Media Bow to Foreign Economic Powers? Evidence from a News Website Crackdown." *Working Paper* .

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

Clayton, Christopher, Matteo Maggiori, and Jesse Schreger. 2023. "A Framework for Geoeconomics." Tech. rep., National Bureau of Economic Research.

———. 2025. "Putting Economics Back Into Geoeconomics." Tech. rep., National Bureau of Economic Research.

Coppedge, Michael et al. 2023. "V-Dem Dataset V13." *Varieties of Democracy (V-Dem) Project* .

Djankov, Simeon, Caralee McLiesh, Tatiana Nenova, and Andrei Shleifer. 2003. "Who Owns the Media?" *The Journal of Law and Economics* 46 (2):341–382.

Dow, Douglas and Amal Karunaratna. 2006. "Developing a Multidimensional Instrument to Measure Psychic Distance Stimuli." *Journal of International Business Studies* 37:578–602.

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.

Durante, Ruben and Ekaterina Zhuravskaya. 2018. "Attack When the World Is Not Watching? US News and the Israeli-Palestinian Conflict." *Journal of Political Economy* 126 (3):1085–1133.

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin. 2009. "Why Resource-Poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103 (4):645–668.

Enikolopov, Ruben, Alexey Makarin, and Maria Petrova. 2020. "Social Media and Protest Participation: Evidence from Russia." *Econometrica* 88 (4):1479–1514.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2011. "Media and Political Persuasion: Evidence from Russia." *American Economic Review* 101 (7):3253–3285.

Fetzer, Thiemo and Prashant Garg. 2025. "Social and Genetic Ties Drive Skewed Cross-Border Media Coverage of Disasters." *arXiv preprint arXiv:2501.07615* .

Fox News. 2022. "Russia Invades Ukraine in Largest European Attack Since WWII." https://www.foxnews.com/world/russian-invades-ukraine-largest-europe-attack-wwii. Accessed May 21, 2025.

Fu, King-Wa. 2023. "Propagandization of Relative Gratification: How Chinese State Media Portray the International Pandemic." *Political Communication* :1–22.

Gehlbach, Scott and Konstantin Sonin. 2014. "Government Control of the Media." *Journal of Public Economics* 118:163–171.

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

Goldsmith, Benjamin E., Yusaku Horiuchi, and Kelly Matush. 2021. "Does Public Diplomacy Sway Foreign Public Opinion? Identifying the Effect of High-level Visits." *American Political Science Review* 115 (4):1342–1357.

Goldstein, Joshua S. 1992. "A Conflict-Cooperation Scale for WEIS Events Data." *Journal of Conflict Resolution* 36 (2):369–385.

Groseclose, Tim and Jeffrey Milyo. 2005. "A Measure of Media Bias." *The Quarterly*

Journal of Economics 120 (4):1191–1237.

Guriev, Sergei, Nikita Melnikov, and Ekaterina Zhuravskaya. 2021. "3G Internet and Confidence in Government." *The Quarterly Journal of Economics* 136 (4):2533–2613.

Guriev, Sergei and Daniel Treisman. 2019. "Informational Autocrats." *Journal of Economic Perspectives* 33 (4):100–127.

———. 2020a. "The Popularity of Authoritarian Leaders: A Cross-National Investigation." *World Politics* 72 (4):601–638.

———. 2020b. "A Theory of Informational Autocracy." *Journal of Public Economics* 186:104158.

Hatte, Sophie, Etienne Madinier, and Ekaterina Zhuravskaya. 2021. "Reading Twitter in the Newsroom: Web 2.0 and Traditional-Media Reporting of Conflicts." *Available at SSRN 3845739* .

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2):994–1028.

Herre, Bastian. 2023. "Identifying Ideologues: A Global Dataset on Political Leaders, 1945–2020." *British Journal of Political Science* 53 (2):740–748.

Holz, Heidi. 2024. "China's Global Public Opinion War with the United States and the West." URL https://warontherocks.com/2024/08/chinas-global-public-opinion-war-with-the-united-states-and-the-west/. War on the Rocks commentary.

Hu, Minqing and Bing Liu. 2004. "Mining and Summarizing Customer Reviews." In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 168–177.

Jervis, Robert. 1976. *Perception and misperception in international politics*. Princeton University Press.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107 (2):326–343.

———. 2017. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument." *American Political Science Review* 111 (3):484–501.

Kleinman, Benny, Ernest Liu, and Stephen J. Redding. 2024. "International Friends and Enemies." *American Economic Journal: Macroeconomics* 16 (4):350–385.

Korovkin, Vasily and Alexey Makarin. 2023. "Conflict and Intergroup Trade: Evidence from the 2014 Russia-Ukraine crisis." *American Economic Review* 113 (1):34–70.

Korovkin, Vasily, Alexey Makarin, and Yuhei Miyauchi. 2025. "Supply Chain Disruption and Reorganization: Theory and Evidence from Ukraine's War." *The Review of Economic Studies* .

Larch, Mario and Yoto V Yotov. 2024. "Estimating the Effects of Trade Agreements: Lessons from 60 years of Methods and Data." *The World Economy* 47 (5):1771–1799.

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.

Leetaru, Kalev and Philip Schrodt. 2013. "Gdelt: Global Data on Events, Location, and Tone, 1979–2012." In *ISA Annual Convention*, vol. 2. Citeseer, 1–49.

Liu, Ernest and David Yang. 2024. "International Power." *CESifo Working Paper Series* .

Lührmann, Anna, Marcus Tannenberg, and Staffan I Lindberg. 2018. "Regimes of the World (RoW): Opening New Avenues for the Comparative Study of Political Regimes." *Politics and Governance* 6 (1):60–77.

Maerz, Seraphine F., Amanda B. Edgell, Matthew C. Wilson, Sebastian Hellmeier, and Staffan I. Lindberg. 2023. "Episodes of Regime Transformation." *Journal of Peace Research* (Online).

Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2019. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2018." *Center for Systemic Peace* 17.

Martin, Gregory J. and Ali Yurukoglu. 2017. "Bias in Cable news: Persuasion and Polarization." *American Economic Review* 107 (9):2565–2599.

Martinez, Luis R. 2022. "How Much Should We Trust the Dictator's GDP Growth Estimates?" *Journal of Political Economy* 130 (10):2731–2769.

McMillan, John and Pablo Zoido. 2002. "How to Subvert Democracy: Montesinos in Peru." *Journal of Economic Perspectives* 18 (4):69–92.

Melitz, Jacques and Farid Toubal. 2014. "Native Language, Spoken Language, Translation and Trade." *Journal of International Economics* 93 (2):351–363.

Mohammad, Saif M. and Peter D. Turney. 2013. "Crowdsourcing a Word–Emotion Association Lexicon." *Computational Intelligence* 29 (3):436–465.

Mohr, Cathrin and Christoph Trebesch. 2024. "Geoeconomics." *CESifo Working Paper Series* .

Nielsen, Finn Årup. 2011. "A New ANEW: Evaluation of a Word List for Sentiment Analysis in Microblogs." *arXiv preprint arXiv:1103.2903* .

Nye, Joseph S. 2004. "Soft power: The means to success in world politics." *Public Affairs* 10.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von Römer. 2022. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." *The Varieties of Democracy Institute* Working Paper.

People's Daily. 2022. "Russia–Ukraine Conflicts Intensify, Sending Shock Waves Across World." http://en.people.cn/n3/2022/0226/c90000-9963162.html. Accessed May 21, 2025.

Petersen, Mitchell A. 2008. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *The Review of Financial Studies* 22 (1):435–480.

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

Puglisi, Riccardo and James Snyder. 2015. "Empirical Studies of Media Bias." In *Handbook of Media Economics*, vol. 1, edited by Simon P. Anderson, Joel Waldfogel, and David Strömberg. Elsevier, 647–667.

Qian, Nancy and David Yanagizawa. 2009. "The Strategic Determinants of U.S. Human Rights Reporting: Evidence from the Cold War." *Journal of the European Economic Association* 7 (2-3):446–457.

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.

Qin, Bei, David Strömberg, and Yanhui Wu. 2018. "Media Bias in China." *American Economic Review* 108 (9):2442–2476.

Roberts, Margaret. 2018. *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton: Princeton University Press.

Rozenas, Arturas and Denis Stukal. 2019. "How Autocrats Manipulate Economic News: Evidence from Russia's State-Controlled Television." *The Journal of Politics* 81 (3):982–996.

Saich, Anthony. 2014. "Reflections on a Survey of Global Perceptions of International Leaders and World Powers." *Ash Center Policy Briefs Series* .

Shaver, Andrew et al. 2022. "News Media Reporting Patterns and Our Biased Understanding of Global Unrest." *Princeton University Empirical Studies of Conflict Working Paper Series* .

Simonov, Andrey and Justin Rao. 2022. "Demand for Online News under Government Control: Evidence from Russia." *Journal of Political Economy* 130 (2):259–309.

Spolaore, Enrico and Romain Wacziarg. 2009. "The Diffusion of Development." *The Quarterly Journal of Economics* 124 (2):469–529.

Strapparava, Carlo and Alessandro Valitutti. 2004. "Word Net-affect: An Affective Extension of WordNet." *Proceedings of the 4th International Conference on Language Resources and Evaluation* .

Strömberg, David. 2015. "Media and Politics." *Annual Review of Economics* 7 (1):173–205.

Sun, Liyang and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of econometrics* 225 (2):175–199.

Tella, Rafael Di and Ignacio Franceschelli. 2011. "Government Advertising and Media Coverage of Corruption Scandals." *American Economic Journal: Applied Economics* 3 (4):119–151.

The Guardian. 2022. "Putin Invades." https://pbs.twimg.com/media/FMZJoxsXoAE6Wsk?format=jpg&name=4096x4096. Front page. Accessed May 21, 2025.

Voeten, Erik. 2000. "Clashes in the Assembly." *International Organization* 54 (2):185–215.

———. 2013. "Data and Analyses of Voting in the United Nations General Assembly." In *Routledge Handbook of International Organization*, edited by Bob Reinalda. New York: Routledge, 54–66.

Voeten, Erik, Anton Strezhnev, and Michael Bailey. 2023. "United Nations General Assembly Voting Data." Harvard Dataverse. URL https://doi.org/10.7910/DVN/LEJUQZ. Accessed: 2023-12-20.

Weiss, Jessica Chen. 2014. *Powerful Patriots: Nationalist Protest in China's Foreign Relations*. Oxford: Oxford University Press.

Xia, Shouzhi, Huang Huang, and Dong Zhang. 2022. "Framing as an Information Control Strategy in Times of Crisis." *Journal of East Asian Studies* 22 (2):255–279.

You, Jiaxing, Bohui Zhang, and Le Zhang. 2018. "Who Captures the Power of the Pen?" *The Review of Financial Studies* 31 (1):43–96.

Young, Lori and Stuart Soroka. 2012. "Affective Eews: The Automated Coding of Sentiment in Political Texts." *Political Communication* 29 (2):205–231.

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov. 2020. "Political Effects of the Internet and Social Media." *Annual Review of Economics* 12:415–438.

Tables

Table 1: Baseline Results

	Dependent Variable: Monthly Average Tone		
	(1)	(2)	(3)
Institutional Distance	-0.280***	-0.263***	-0.260***
	(0.046)	(0.045)	(0.043)
Observations	81,288,434	81,288,434	81,288,434
Number of Country Pairs	14,379	14,379	14,379
Target Country × Theme × Year FEs	Yes	Yes	Yes
Outlet FEs	Yes	No	No
Outlet × Year FEs	No	Yes	No
Outlet × Theme × Year FEs	No	No	Yes

Notes: This table documents the presence of systematic bias in the tone with which media outlets cover institutionally distant countries. The dependent variable is the monthly average tone from GDELT GKG 2.0 database measured at the media-outlet–target-country–theme–month level. Other variables are defined in Online Appendix F. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 2: Biased News Coverage of National Leaders

	Dependent Variable: Tone		
	(1)	(2)	(3)
Institutional Distance	-0.969***	-0.916***	-0.859***
	(0.163)	(0.142)	(0.123)
Std of Tone	3.418	3.418	3.418
Std of Institutional Distance	0.227	0.227	0.227
Observations	426,658	424,558	408,696
Number of Leaders	339	339	339
Number of Country Pairs	3,262	3,218	3,005
Outlet FEs	Yes	No	No
Leader FEs	Yes	No	No
Year FEs	Yes	No	No
Outlet × Year FEs	No	Yes	No
Leader × Year FEs	No	Yes	No
Outlet × Year-Month FEs	No	No	Yes
Leader × Year-Month FEs	No	No	Yes

Notes: This table examines the relationship between institutional distance and media bias toward national leaders. The sample is constructed at the news-article level and includes all articles in which a unique national leader is identified through a systematic keyword search (see Online Appendix B.2). Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 3: Regime Types and Information Control

Panel A: Heterogeneity by Regime Types

	Dependent Variable: Monthly Average Tone			
	(1)	(2)	(3)	(4)
Institutional Distance	-0.394***	-0.387***	-0.119*	-0.106*
	(0.064)	(0.062)	(0.064)	(0.062)
Institutional Distance × Source Country Is a Liberal Democracy	0.251***	0.246***		
	(0.080)	(0.078)		
Institutional Distance × Target Country Is a Liberal Democracy			-0.298***	-0.318***
			(0.115)	(0.111)
Observations	81,288,434	81,288,434	81,288,434	81,288,434
Number of Country Pairs	14,379	14,379	14,379	14,379

Panel B: Heterogeneity by Media Freedom

	Dependent Variable: Monthly Average Tone			
	(1)	(2)	(3)	(4)
Institutional Distance	-0.352***	-0.346***	-0.329***	-0.327***
	(0.060)	(0.058)	(0.057)	(0.055)
Institutional Distance × Freedom of Traditional Media	0.097***	0.095***		
	(0.027)	(0.026)		
Institutional Distance × Freedom of the Internet			0.117***	0.119***
			(0.040)	(0.039)
Observations	81,288,434	81,288,434	81,288,434	81,288,434
Number of Country Pairs	14,379	14,379	14,379	14,379

Panel C: Heterogeneity by Media Ownership

	Dependent Variable: Monthly Average Tone			
	(1)	(2)	(3)	(4)
Institutional Distance	-0.433***	-0.423***	-0.329***	-0.322***
	(0.084)	(0.081)	(0.058)	(0.057)
Institutional Distance × State-owned Print Media	-0.145***	-0.140***		
	(0.045)	(0.043)		
Institutional Distance × State-owned Broadcast Media			-0.076**	-0.072**
			(0.034)	(0.032)
Observations	79,889,616	79,889,616	79,855,427	79,855,427
Number of Country Pairs	13,838	13,838	13,839	13,839
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes
Outlet × Year FEs	Yes	No	Yes	No
Outlet × Theme × Year FEs	No	Yes	No	Yes

Notes: This table examines heterogeneity in the association between institutional distance and media sentiment across regime characteristics. Panel A examines whether the magnitude of the institutional distance coefficient varies with the regime types of the source and target countries. Panel B examines whether the coefficient varies with the level of media freedom in the source country. Panel C examines whether the coefficient varies with the degree of state ownership of media in the source country. Other model specifications follow the conventions in Table 1. All variables are defined in Online Appendix F. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Economic Distress

	Dependent Variable: Monthly Average Tone			
	(1)	(2)	(3)	(4)
Institutional Distance	-0.248***	-0.244***	-0.236***	-0.233***
	(0.047)	(0.045)	(0.050)	(0.048)
Monthly Growth in Nighttime Light Intensity	-0.006	-0.006		
	(0.010)	(0.010)		
Institutional Distance × Monthly Growth in Nighttime Light Intensity	0.095***	0.090**		
	(0.036)	(0.036)		
Institutional Distance × High Inflation Indicator			-0.155*	-0.154*
			(0.086)	(0.083)
Observations	66,414,057	66,411,598	80,117,942	80,117,942
Number of Country Pairs	12,640	12,640	14,326	14,326
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes
Outlet × Year FEs	Yes	No	Yes	No
Outlet × Theme × Year FEs	No	Yes	No	Yes

Notes: This table examines heterogeneity in the association between institutional distance and media sentiment by the source country's economic conditions. Other model specifications follow the conventions in Table 1. All variables are defined in Online Appendix F. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 5: Domestic Turmoil

	All Source Countries		Not Liberal Democracies	Liberal Democracies
	Dependent Variable: Monthly Average Tone			
	(1)	(2)	(3)	(4)
Institutional Distance	-0.260***	-0.237***	-0.243**	-0.346*
	(0.043)	(0.084)	(0.099)	(0.201)
Domestic Turmoil Index	-0.602***	-0.543**	-0.108	-0.661**
	(0.094)	(0.220)	(0.257)	(0.303)
Institutional Distance × Domestic Turmoil Index		-0.201	-1.722**	0.552
		(0.706)	(0.761)	(0.996)
Observations	81,288,434	81,288,434	32,836,225	48,451,923
Number of Country Pairs	14,379	14,379	13,521	6,217
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes
Outlet × Theme × Year FEs	Yes	Yes	Yes	Yes

Notes: This table examines heterogeneity in the association between institutional distance and media sentiment by the level of domestic turmoil in the source country. Columns (3) and (4) restrict the sample to source countries that are not liberal democracies and to liberal democracies, respectively. Other model specifications follow the conventions in Table 1. All variables are defined in Online Appendix F. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 6: Biased Theme Selection

	Dependent Variable				
	Share		Report (Binary)		Share (Violence)
	(1)	(2)	(3)	(4)	(5)
Institutional Distance	-0.834***	-0.391***	-0.120***	-0.122***	1.428***
	(0.048)	(0.032)	(0.007)	(0.007)	(0.159)
Institutional Distance × Tone (Theme)	-0.414***		-0.012***		
	(0.031)		(0.001)		
Tone (Target-Theme-Year)		-0.160***		-0.012***	
		(0.006)		(0.001)	
Institutional Distance × Tone (Target-Theme-Year)		-0.162***		-0.016***	
		(0.016)		(0.002)	
Std of Dependent Variables	7.008	7.008	0.481	0.481	13.758
Std of Institutional Distance	0.212	0.212	0.212	0.212	0.212
Std of Tone (Theme)	0.863	0.863	0.863	0.863	—
Std of Tone (Target-Theme-Year)	1.546	1.546	1.546	1.546	—
Observations	5,947,392	5,925,850	5,947,392	5,925,850	247,808
Number of Country Pairs	15,576	15,576	15,576	15,576	15,576
Target × Theme FE	Yes	Yes	Yes	Yes	Yes
Source Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

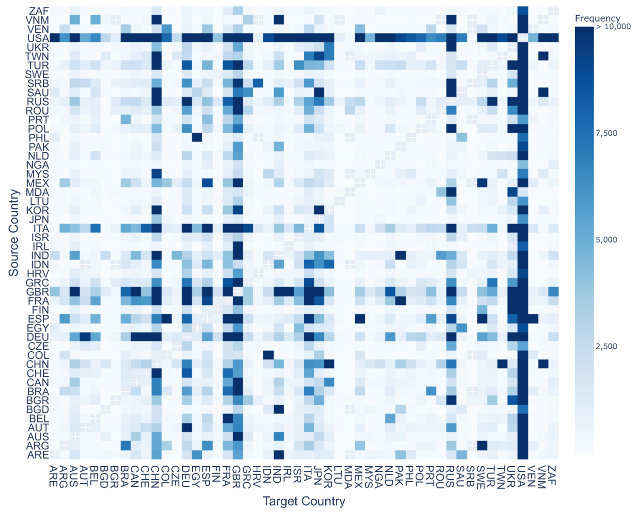
Notes: This table examines the presence of global media bias from the perspective of theme selection. We construct the sample at the source-country–target-country–theme–year level with information from the baseline sample. *Tone (Theme)* is defined as the average tone across all articles classified under a given theme. *Tone (Target–Theme–Year)* is defined as the average tone of all articles covering a given target country on a particular theme in a given year, excluding articles published by outlets from the source country under consideration. The dependent variable in columns (1)–(2) is the share of articles in a given theme, scaled by the total number of articles published by the source country about the target country in that year. The dependent variable in columns (3)–(4) is an indicator for whether a given theme is covered at all by the source country in its coverage of the target country. The dependent variable in column (5) is the share of articles classified under the theme "Fragility, Conflict, and Violence." Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 7: Public Opinion Reactions

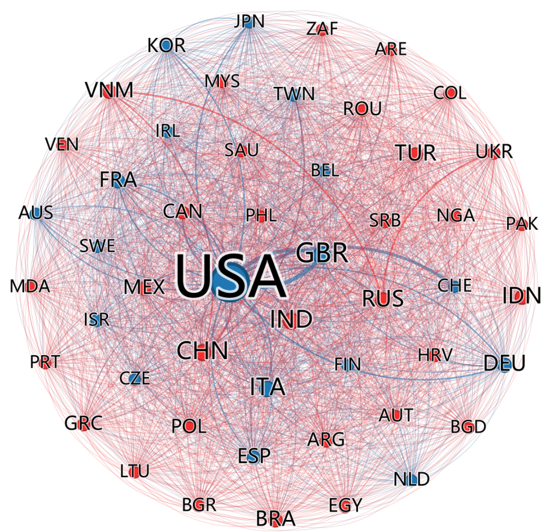
	Dependent Variable: Approval		
	(1)	(2)	(3)
Tone (Previous Year)	0.023***	0.008***	0.009**
	(0.005)	(0.003)	(0.004)
Mean of Approval	0.521	0.521	0.521
Std of Approval	0.500	0.500	0.500
Mean of Tone	-1.421	-1.421	-1.421
Std of Tone	1.263	1.263	1.263
Observations	3,570,351	3,570,351	3,570,351
Number of Source Countries	148	148	148
Number of Target Countries	9	9	9
Number of Country Pairs	786	786	786
Control Variables	Yes	Yes	Yes
Target Country FEs	Yes	Yes	No
Year FEs	Yes	Yes	No
Source Country FEs	No	Yes	No
Source Country × Year FEs	No	No	Yes
Target Country × Year FEs	No	No	Yes

Notes: This table presents evidence on whether the tone of global media coverage predicts subsequent public opinion of foreign countries. Public opinion data is sourced from the Gallup World Poll. The main independent variable is the average tone of news articles from a source country toward the target country in the previous year. The dependent variable, *Approval*, equals one if the respondent approves of the leadership performance of the target country, and zero otherwise. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Figures



(a) Bilateral Frequency



(b) Reporting Network

Figure 1: Bilateral Reporting Relations

Notes: Panel A depicts a heat map of cross-country media coverage intensity for the top 50 countries by number of news articles in 2022. The y-axis is the source country; the x-axis is the target country. Panel B illustrates the network structure with the same data used in Panel A. Node size indicates total media articles; edge thickness shows coverage volume, and blue color indicates liberal democracies.

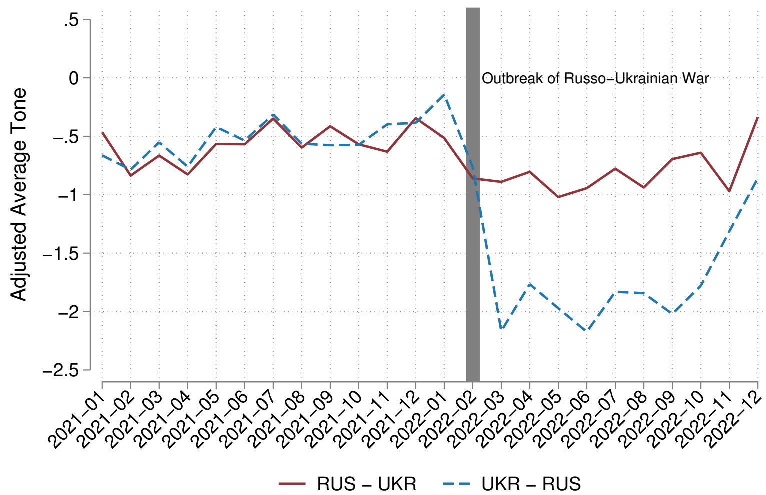


Figure 2: Changes in Media Tone Between Russia and Ukraine

Notes: This figure validates the tone measure by plotting adjusted media tone in Russian and Ukrainian coverage before and after Russia's invasion of Ukraine in February 2022. The solid red line shows Russian media sentiment toward Ukraine, and the dashed blue line shows Ukrainian media sentiment toward Russia. Tone is residualized with respect to theme-by-month fixed effects.

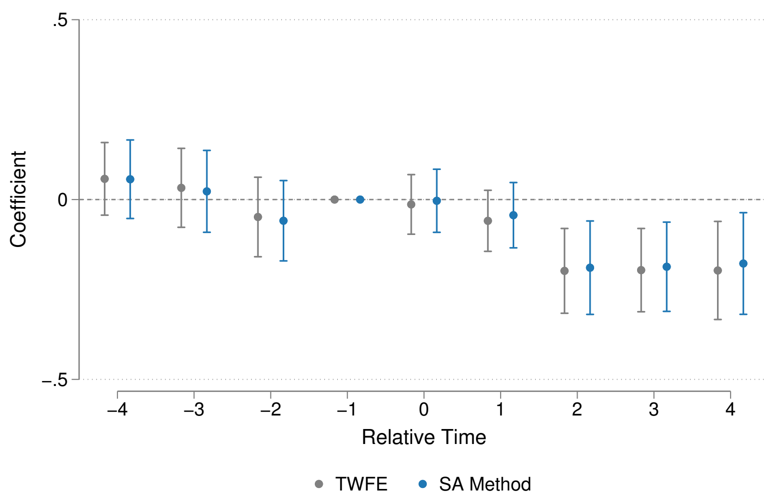
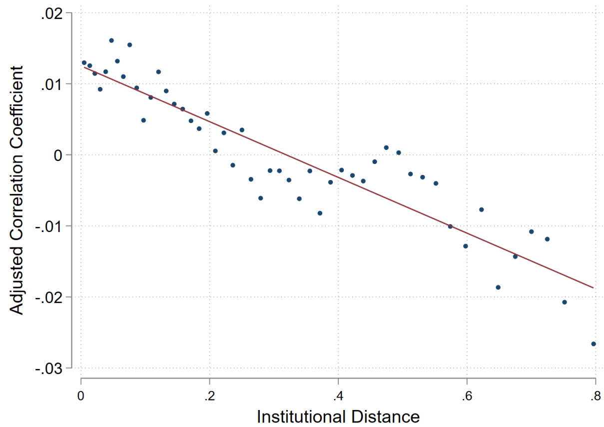
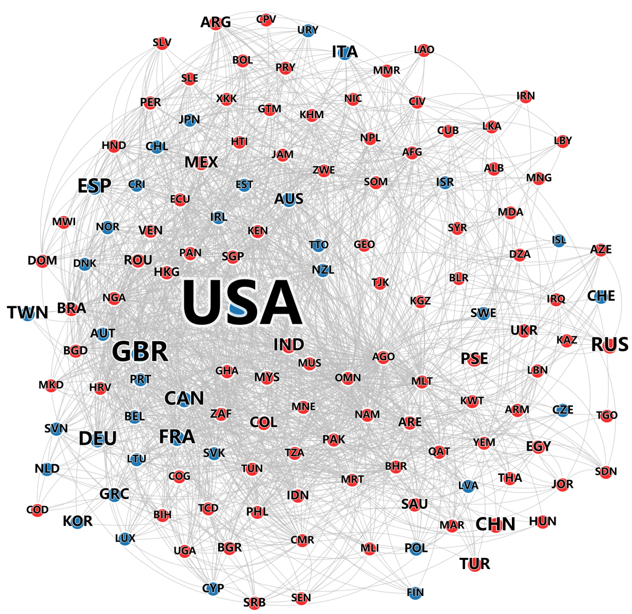


Figure 3: Decline in Average Coverage Tone Toward Liberal Democracies After Democratic Erosion in the Source Country

Notes: This figure plots coefficient estimates from Equation (6), studying changes in coverage tone toward liberal democracies following democratic erosion in source countries. Estimates are shown for both the TWFE and the Sun and Abraham (2021) specifications. Bars represent the 95% confidence intervals. See Table A19 for the full set of estimates.



(a) Institutional Distance and Alliances



(b) Alliance Networks

Figure 4: Institutional Distance and Alliance Networks

Notes: This figure illustrates media alliances in the tone of countries' coverage of *third* countries. Panel A plots binned averages of pairwise media-tone correlation coefficients (ρ), computed from countries' coverage of other countries, against bilateral institutional distance. The correlation coefficients (ρ) are residualized with respect to target-country, source-country, and year fixed effects. Panel B shows a network representation of media alliances. Nodes denote countries, and an edge links two countries if their correlation coefficient over the 2015–2022 period falls in the top quintile. Blue nodes indicate liberal democracies and red nodes indicate nonliberal democracies. Node size is proportional to the volume of news articles produced by outlets in that country.

Online Appendix

Table of Contents

A Additional Tables and Figures

Table A1: Summary Statistics

Panel A: Variables at the Media-Outlet–Target-Country–Theme–Month Level (Baseline Sample)

	N	Mean	Median	Std	P25	P75
Monthly Average Tone	81,288,434	-1.929	-1.677	3.747	-4.228	0.581
Monthly Average Positive Tone	81,288,434	2.722	2.466	1.663	1.589	3.548
Monthly Average Negative Tone	81,288,434	4.650	4.274	2.889	2.483	6.343

Panel B: Variables at the Country-Pair–Year Level

	N	Mean	Median	Std	P25	P75
Institutional Distance	97,340	0.313	0.276	0.223	0.119	0.482
Institutional Distance (Electoral Democracy)	97,340	0.303	0.267	0.214	0.119	0.463
Institutional Distance (Polity Index)	56,733	6.577	5.000	5.685	2.000	11.000
UN Voting Distance	91,227	0.993	0.844	0.787	0.298	1.561
Common Language	89,426	0.140	0.000	0.347	0.000	0.000
Religious Distance	75,857	5.085	4.252	2.971	2.500	8.247
Geographic Distance	89,426	7181.598	6604.113	4323.596	3758.281	10007.979
Genetic Distance	82,821	0.033	0.036	0.017	0.016	0.046
Trade Dependence	83,636	0.006	0.000	0.019	0.000	0.003
Trade Balance	83,553	-0.001	0.000	0.009	0.000	0.000
Sanctioned	74,373	0.017	0.000	0.105	0.000	0.000

Panel C: Variables at the Country–Year or Country–Year-Month Level

	N	Mean	Median	Std	P25	P75
Source Country						
Liberal Democracy Index	1,385	0.411	0.384	0.264	0.162	0.640
Liberal Democracy Indicator (Yes=1)	1,385	0.222	0.000	0.416	0.000	0.000
Monthly Growth in Nighttime Light Intensity	11,230	0.007	0.002	0.380	-0.118	0.122
High Inflation Indicator	1,343	0.302	0.000	0.459	0.000	1.000
Freedom of Traditional Media	1,385	0.613	0.796	1.537	-0.427	1.723
Freedom of the Internet	1,385	0.542	0.941	1.377	-0.453	1.733
State-owned Print Media	1,220	-0.806	-0.836	1.236	-1.718	0.028
State-owned Broadcast Media	1,220	-0.624	-0.858	0.987	-1.247	-0.120
Domestic Turmoil Index	15,963	0.100	0.100	0.050	0.070	0.130
Target Country						
Liberal Democracy Index	1,416	0.406	0.380	0.264	0.155	0.637
Liberal Democracy Indicator (Yes=1)	1,416	0.217	0.000	0.412	0.000	0.000

Notes: Panel A reports the main dependent variables at the media-outlet–target-country–theme–month level. Panel B reports the independent variables at the country-pair level. Panel C reports country–year- and country–year-month variables primarily used to examine mechanisms underlying our baseline results. Sample construction is detailed in Online Appendix B. The definitions of all variables are provided in Online Appendix F.

Table A2: Additional Summary Statistics

Panel A: Summary Statistics by Selected Country Pairs

	Tone				
Source Country - Target Country	Mean	Median	Std	P25	P75
USA - GBR	-0.800	-0.633	3.245	-2.710	1.297
GBR - USA	-0.896	-0.858	3.159	-3.107	1.331
USA - CHN	-1.551	-1.408	3.235	-3.548	0.571
CHN - USA	-1.379	-1.237	3.019	-3.090	0.398
USA - RUS	-2.546	-2.400	3.387	-4.706	-0.284
RUS - USA	-1.912	-1.887	2.840	-3.571	-0.183
CHN - RUS	-0.424	-0.263	3.564	-2.463	1.823
RUS - CHN	-1.467	-1.306	3.022	-3.246	0.469
USA - IRN	-3.079	-2.825	3.660	-5.405	-0.502
IRN - USA	-2.833	-2.859	3.019	-4.782	-0.879

Panel B: Summary Statistics by Top Ten Themes

	Tone				
Theme	Mean	Median	Std	P25	P75
Public Sector Management	-2.579	-2.473	3.785	-4.952	0.000
Health Nutrition and Population	-2.470	-2.290	3.785	-4.764	0.000
Fragility Conflict and Violence	-3.795	-3.682	3.914	-6.291	-1.158
Information and Communication Technologies	-1.762	-1.508	3.751	-4.104	0.771
Education	-1.405	-1.205	3.681	-3.596	0.991
Private Sector Development	-1.624	-1.401	3.622	-3.791	0.775
Transport	-2.037	-1.812	3.754	-4.444	0.575
Jobs	-0.740	-0.563	3.377	-2.722	1.440
Water	-1.652	-1.393	3.600	-3.846	0.720
Financial Sector Development	-1.633	-1.362	3.377	-3.554	0.544

Table A3: Positive Tone vs. Negative Tone

	Dependent Variable			
	Monthly Average Positive Tone		Monthly Average Negative Tone	
	(1)	(2)	(3)	(4)
Institutional Distance	-0.050***	-0.049***	0.213***	0.211***
	(0.017)	(0.017)	(0.034)	(0.032)
Observations	81,288,434	81,288,434	81,288,434	81,288,434
Number of Country Pairs	14,379	14,379	14,379	14,379
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes
Outlet × Year FEs	Yes	No	Yes	No
Outlet × Theme × Year FEs	No	Yes	No	Yes

Notes: This table presents the baseline estimates of the association between institutional distance and monthly positive and negative tones. *Positive Tone* and *Negative Tone*, denote, respectively, the percentage of positive and negative words in news articles. We aggregate tone measures at the outlet–target-country–theme–month level. Other model specifications follow the convention in Table 1. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A4: Alternative Measure of Geopolitical Alignment

Panel A. UN Voting Distance

	Dependent Variable: Monthly Average Tone		
	(1)	(2)	(3)
UN Voting Distance	-0.059***	-0.054***	-0.054***
	(0.012)	(0.012)	(0.011)
Observations	78,781,874	78,781,645	78,774,714
Number of Country Pairs	13,501	13,501	13,501

Panel B. All Distance Measures

	Dependent Variable: Monthly Average Tone		
	(1)	(2)	(3)
Institutional Distance	-0.232***	-0.229***	-0.221***
	(0.058)	(0.057)	(0.055)
UN Voting Distance	0.015	0.019	0.016
	(0.015)	(0.016)	(0.014)
Common Language	0.229***	0.227***	0.219***
	(0.029)	(0.028)	(0.027)
Religious Distance	-0.001	-0.001	-0.002
	(0.006)	(0.006)	(0.006)
Geographic Distance	-0.020	-0.019	-0.016
	(0.014)	(0.013)	(0.013)
Genetic Distance	-0.568	-0.557	-0.420
	(1.474)	(1.426)	(1.357)
Observations	69,187,422	69,185,980	69,133,736
Number of Country Pairs	8,944	8,944	8,944
Std of Monthly Average Tone	3.747	3.747	3.747
Std of Institutional Distance	0.223	0.223	0.223
Std of UN Voting Distance	0.787	0.787	0.787
Target Country × Theme × Year FEs	Yes	Yes	Yes
Outlet FEs	Yes	No	No
Outlet × Year FEs	No	Yes	No
Outlet × Theme × Year FEs	No	No	Yes

Notes: This table examines media bias in coverage of geopolitically distant countries. Panel A reports results using UN voting distance as the key independent variable. Panel B reports results including all distance measures, including our primary independent variable of institutional distance. Variables are defined in Online Appendix F. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A5: Baseline Estimates by Year

	Dependent Variable: Monthly Average Tone			
	(1)	(2)	(3)	(4)
Institutional Distance	-0.193***	-0.191***		
	(0.045)	(0.043)		
Linear Time Trend × Institutional Distance	-0.024***	-0.023***		
	(0.008)	(0.008)		
D_{2015} × Institutional Distance			-0.208***	-0.205***
			(0.048)	(0.047)
D_{2016} × Institutional Distance			-0.217***	-0.217***
			(0.047)	(0.046)
D_{2017} × Institutional Distance			-0.221***	-0.224***
			(0.046)	(0.045)
D_{2018} × Institutional Distance			-0.265***	-0.261***
			(0.051)	(0.050)
D_{2019} × Institutional Distance			-0.292***	-0.280***
			(0.058)	(0.056)
D_{2020} × Institutional Distance			-0.294***	-0.282***
			(0.055)	(0.053)
D_{2021} × Institutional Distance			-0.375***	-0.375***
			(0.062)	(0.060)
D_{2022} × Institutional Distance			-0.352***	-0.350***
			(0.068)	(0.067)
Observations	81,288,433	81,288,434	81,288,432	81,288,434
Number of Country Pairs	14,379	14,379	14,379	14,379
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes
Outlet × Year FEs	Yes	No	Yes	No
Outlet × Theme × Year FEs	No	Yes	No	Yes

Notes: This table presents the year-by-year association between institutional distance and media sentiment. D_t is an indicator variable that is set to one for year t. Other model specifications follow the conventions in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A6: Alternative Model Specifications

	Dependent Variable: Monthly Average Tone				
	(1)	(2)	(3)	(4)	(5)
Institutional Distance	-0.158**	-0.156**	-0.247***	-0.166***	-0.160***
	(0.063)	(0.062)	(0.043)	(0.058)	(0.057)
Observations	81,288,234	81,287,826	74,700,750	74,700,489	74,699,889
Number of Country Pairs	14,185	14,172	14,340	14,120	14,101
Target Country × Theme × Year FEs	Yes	Yes	No	No	No
Outlet × Theme × Year FEs	Yes	Yes	No	No	No
Target Country × Theme × Month FEs	No	No	Yes	Yes	Yes
Outlet × Theme × Month FEs	No	No	Yes	Yes	Yes
Country Pair FEs	Yes	No	No	Yes	No
Country Pair (Directed) FEs	No	Yes	No	No	Yes

Notes: We consider alternative model specifications with more restrictive fixed effects in addition to the baseline models in Table 1. Other variable construction and model specifications follow the convention in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A7: Alternative Measures of Tone

	Dependent Variable			
	Monthly Average Tone (LSD)		Monthly Average Tone (SWN)	
	(1)	(2)	(3)	(4)
Institutional Distance	-0.176***	-0.175***	-0.001***	-0.002***
	(0.045)	(0.043)	(0.000)	(0.000)
Observations	81,288,434	81,288,434	81,288,433	81,288,433
Number of Country Pairs	14,379	14,379	14,379	14,379
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes
Outlet × Year FEs	Yes	No	Yes	No
Outlet × Theme × Year FEs	No	Yes	No	Yes

Notes: This table presents results based on alternative sentiment measures. We consider two alternatives: the Lexicoder Sentiment Dictionary (LSD) and SentiWordNet 3.0 (SWN). Other variable construction and model specifications follow the convention in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A8: Alternative Measures of Institutional Distance

	Dependent Variable: Monthly Average Tone			
	(1)	(2)	(3)	(4)
Institutional Distance (Electoral Democracy)	-0.304***	-0.300***		
	(0.051)	(0.049)		
Institutional Distance (Polity Index)			-0.014***	-0.014***
			(0.003)	(0.003)
Observations	81,288,434	81,288,434	57,984,858	57,980,165
Number of Country Pairs	14,379	14,379	12,684	12,684
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes
Outlet × Year FEs	Yes	No	Yes	No
Outlet × Theme × Year FEs	No	Yes	No	Yes

Notes: This table presents robustness checks to the baseline specification using alternative measures of institutional distance. We consider two alternatives: the V-Dem Electoral Democracy Index and the Polity Index. Other variable definitions and model specifications follow the conventions in Table 1. Further details on variable construction are provided in Online Appendix F. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A9: Robustness Checks on Translation-Related Measurement

	Dependent Variable			
	Tone	Monthly Average Tone		
	(1)	(2)	(3)	(4)
Translation	0.004			
	(0.018)			
Institutional Distance		-0.260***	-0.277***	-0.299***
		(0.043)	(0.051)	(0.054)
Pct(English Articles)		-0.099***	-0.116***	
		(0.025)	(0.032)	
Institutional Distance × Pct(English Articles)			0.052	
			(0.066)	
Observations	4,703,783	81,288,434	81,288,434	53,809,122
Number of Country Pairs	-	14,379	14,379	9,822
Outlet × Year FEs	Yes	No	No	No
Target × Year FEs	Yes	No	No	No
Theme FEs	Yes	No	No	No
Target Country × Theme × Year FEs	No	Yes	Yes	Yes
Outlet × Theme × Year FEs	No	Yes	Yes	Yes

Notes: This table presents robustness checks on the baseline results, addressing potential measurement issues related to translation. Column (1) uses a random sample selecting 50,000 articles per month from the GDELT GKG 2.0 database to estimate the correlation between a translation dummy variable (equal to one if the article is machine-translated) and article-level tone. In columns (2) and (3), we construct a variable capturing the proportion of English-language articles in the baseline sample and examine the correlation between institutional distance and media tone, controlling for the proportion of English-language articles in column (2) and additionally controlling for its interaction with institutional distance in column (3). Column (4) restricts the sample to countries where one of the world's ten most widely spoken languages is official, addressing concerns that the results may be driven by translation quality in low-resource language environments. Other variable construction and model specifications follow the convention in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A10: Domain-Source Determination Rules

	Dependent Variable: Monthly Average Tone				
	(1)	(2)	(3)	(4)	(5)
Institutional Distance	-0.263***	-0.240***	-0.262***	-0.575***	-0.313***
	(0.042)	(0.062)	(0.043)	(0.066)	(0.059)
Domain Source Rule	Most Frequency	DNS	GPT-4o	Wiki	TLD
Observations	81,198,262	78,975,513	81,202,902	1,693,339	35,135,302
Number of Country Pairs	14,256	10,768	14,351	6,787	12,944
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes	Yes
Outlet × Theme × Year FEs	Yes	Yes	Yes	Yes	Yes

Notes: This table tests robustness of the baseline specification to alternative domain-source determination strategies. The *Most Frequency* rule assigns an outlet to the country it covers most frequently. *DNS* assigns countries to outlets using reverse DNS lookup. *GPT-4o* assigns countries to outlets using classifications from a generative AI model. *Wiki* restricts the sample to a Wikipedia list of 187 major news agencies by country. *TLD* assigns country affiliation based on an outlet's top-level domain suffix (e.g., .uk, .ru). Other variable constructions and model specifications follow the conventions in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A11: Winsorization

	Dependent Variable: Monthly Average Tone		
	Winsor 0.5%	Winsor 1%	Winsor 5%
	(1)	(2)	(3)
Institutional Distance	-0.264***	-0.266***	-0.264***
	(0.043)	(0.042)	(0.039)
Observations	81,288,433	81,288,433	81,288,432
Number of Country Pairs	14,379	14,379	14,379
Target Country × Theme × Year FEs	Yes	Yes	Yes
Outlet × Theme × Year FEs	Yes	Yes	Yes

Notes: This table tests robustness of the baseline specification to winsorizing the dependent variable, *Monthly Average Tone*. We winsorize the dependent variable at 0.5%, 1%, and 5%, respectively. Other variable construction and model specifications follow the convention in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A12: Excluding Observations Related to the United States

	Dependent Variable: Monthly Average Tone		
	(1)	(2)	(3)
Institutional Distance	-0.296***	-0.276***	-0.303***
	(0.044)	(0.045)	(0.047)
Observations	63,127,811	76,624,788	56,279,085
Number of Country Pairs	14,265	14,265	14,089
Drop USA as Source Country	Yes	No	Yes
Drop USA as Target Country	No	Yes	Yes
Target Country × Theme × Month FEs	Yes	Yes	Yes
Outlet × Theme × Year FEs	Yes	Yes	Yes

Notes: This table tests robustness of the baseline specification to removing observations related to the United States. Variable construction and other model specifications follow the convention in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A13: Alternative Sample-Level

	Dependent Variable: Monthly Average Tone	
	(1)	(2)
Institutional Distance	-0.226***	-0.223***
	(0.031)	(0.031)
Observations	11,006,336	11,006,336
Number of Country Pairs	14,379	14,379
Target Country × Theme × Year FEs	Yes	Yes
Source Country × Year FEs	Yes	No
Source Country × Theme × Year FEs	No	Yes

Notes: This table tests robustness of the baseline specification to collapsing the baseline sample to the country-pair–theme–year level. Other variable construction and model specifications follow the convention in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A14: Alternative Sample Construction

	Dependent Variable: Monthly Average Tone		
	(1)	(2)	(3)
Institutional Distance	-0.350***	-0.332***	-0.355***
	(0.035)	(0.034)	(0.032)
Std of Monthly Average Tone	3.988	3.988	3.988
Std of Institutional Distance	0.226	0.226	0.226
Observations	34,958,288	34,958,288	34,958,288
Number of Country Pairs	14,200	14,200	14,200
Target Country × Event-Type × Year FEs	Yes	Yes	Yes
Outlet FEs	Yes	No	No
Outlet × Year FEs	No	Yes	No
Outlet × Event-Type × Year FEs	No	No	Yes

Notes: We replicate the baseline results in Table 1 using data from the GDELT Event 2.0 database. The sample is at the media-outlet–target-country–event-type–month level. Details on the construction of the GDELT Event 2.0 sample are provided in Online Appendix B.3. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A15: Controlling for Economic Interdependence and Sanctions

	Dependent Variable: Monthly Average Tone		
	(1)	(2)	(3)
Institutional Distance	-0.253***	-0.255***	-0.250***
	(0.047)	(0.046)	(0.042)
Trade Dependence	0.137		
	(0.092)		
Trade Balance		-0.285	
		(0.277)	
Sanctioned			-0.042
			(0.045)
Observations	77,527,780	77,513,860	81,288,434
Number of Country Pairs	12,863	12,862	14,379
Target Country × Theme × Year FEs	Yes	Yes	Yes
Outlet × Theme × Year FEs	Yes	Yes	Yes

Notes: This table tests robustness of the baseline specification to controlling for trade relationships. Definitions and data sources for these variables are provided in Online Appendix F. Other variable construction and model specifications follow the convention in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A16: Alternative SE Clustering Methods

	Dependent Variable: Monthly Average Tone				
	(1)	(2)	(3)	(4)	(5)
Institutional Distance	-0.260***	-0.260***	-0.260***	-0.260***	-0.260***
	(0.046)	(0.058)	(0.059)	(0.013)	(0.023)
SE Cluster Level	Country-pair & year (two-way)	Source country	Source country & year (two-way)	Outlet	Outlet & year (two-way)
Observations	81,288,434	81,288,434	81,288,434	81,288,434	81,288,434
Number of Clusters 1	14,379	176	176	55,240	55,240
Number of Clusters 2	8	0	8	0	8
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes	Yes
Outlet × Theme × Year FEs	Yes	Yes	Yes	Yes	Yes

Notes: This table presents robustness checks to the baseline estimation using alternative clustering methods for calculating standard errors. Other variable construction and model specifications follow the convention in Table 1. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A17: Timing of the Democratic Erosion

Country	Treatment Year	Change in Regime-Type Classification
Afghanistan	2016	Electoral Autocracy → Closed Autocracy
Albania	2018	Electoral Democracy → Electoral Autocracy
Austria	2021	Liberal Democracy → Electoral Democracy
El Salvador	2018	Electoral Democracy → Electoral Autocracy
Guatemala	2018	Electoral Democracy → Electoral Autocracy
Guinea	2019	Electoral Autocracy → Closed Autocracy
Iran	2021	Electoral Autocracy → Closed Autocracy
Lithuania	2016	Liberal Democracy → Electoral Democracy
Mali	2017	Electoral Democracy → Electoral Autocracy
Myanmar (Burma)	2021	Electoral Autocracy → Closed Autocracy
Nigeria	2021	Electoral Democracy → Electoral Autocracy
Philippines	2016	Electoral Democracy → Electoral Autocracy
Poland	2016	Liberal Democracy → Electoral Democracy
Portugal	2021	Liberal Democracy → Electoral Democracy
Sudan	2019	Electoral Autocracy → Closed Autocracy
Togo	2016	Electoral Democracy → Electoral Autocracy
Trinidad and Tobago	2021	Liberal Democracy → Electoral Democracy

Notes: This table presents the timing of democratic erosions in our sample period, defined as changes in regime category in the V-Dem Regimes of the World (RoW) classification. Details on the construction of this event sample are provided in Section 3.4.

Table A18: Democratic Erosion and Media Bias

	Dependent Variable: Monthly Average Tone	
	(1)	(2)
Democratic Erosion × Liberal Democracy Index (Target Country)	-0.127**	
	(0.060)	
Democratic Erosion × Liberal Democracy Indicator (Target Country)		-0.089***
		(0.032)
Observations	60,330,253	60,330,253
Number of Country Pairs	11,040	11,040
Target Country × Theme × Year FEs	Yes	Yes
Outlet × Theme × Year FEs	Yes	Yes
Outlet × Target Country FEs	Yes	Yes

Notes: This table presents the nondynamic staggered difference-in-differences estimates of whether democratic erosion in the source country leads to more negative coverage of liberal democracies in its outlets. We exclude the always-treated group and source countries classified as closed autocracies, which cannot experience democratic erosion, from the baseline sample. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A19: Democratic Erosion and Media Bias: Dynamic Estimates

	Dependent Variable: Monthly Average Tone		
	TWFE	TWFE	SA Method
	(1)	(2)	(3)
t-4 × Liberal Democracy Index (Target Country)	0.084		
	(0.086)		
t-3 × Liberal Democracy Index (Target Country)	0.059		
	(0.097)		
t-2 × Liberal Democracy Index (Target Country)	-0.064		
	(0.100)		
t × Liberal Democracy Index (Target Country)	0.007		
	(0.069)		
t+1 × Liberal Democracy Index (Target Country)	-0.089		
	(0.080)		
t+2 × Liberal Democracy Index (Target Country)	-0.385***		
	(0.109)		
t+3 × Liberal Democracy Index (Target Country)	-0.321***		
	(0.104)		
t+4 × Liberal Democracy Index (Target Country)	-0.325**		
	(0.133)		
t-4 × Liberal Democracy Indicator (Target Country)		0.058	0.056
		(0.051)	(0.056)
t-3 × Liberal Democracy Indicator (Target Country)		0.033	0.023
		(0.056)	(0.058)
t-2 × Liberal Democracy Indicator (Target Country)		-0.048	-0.059
		(0.056)	(0.057)
t × Liberal Democracy Indicator (Target Country)		-0.013	-0.003
		(0.042)	(0.045)
t+1 × Liberal Democracy Indicator (Target Country)		-0.059	-0.043
		(0.043)	(0.046)
t+2 × Liberal Democracy Indicator (Target Country)		-0.198***	-0.189***
		(0.060)	(0.066)
t+3 × Liberal Democracy Indicator (Target Country)		-0.196***	-0.187***
		(0.059)	(0.063)
t+4 × Liberal Democracy Indicator (Target Country)		-0.197***	-0.178***
		(0.070)	(0.072)
Observations	60,330,253	60,330,253	60,330,253
Number of Country Pairs	11,040	11,040	11,040
Target Country × Theme × Year FEs	Yes	Yes	Yes
Outlet × Theme × Year FEs	Yes	Yes	Yes
Outlet × Target Country FEs	Yes	Yes	Yes

Notes: This table presents the dynamic staggered difference-in-differences estimates of whether democratic erosion in the source country leads to more negative coverage of liberal democracies in its outlets. We exclude the always-treated group and source countries classified as closed autocracies, which cannot experience democratic erosion, from the baseline sample. Columns (1)–(2) use two-way fixed-effects OLS, while column (3) applies the method of Sun and Abraham (2021). Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A20: State-Owned Media and Media Bias

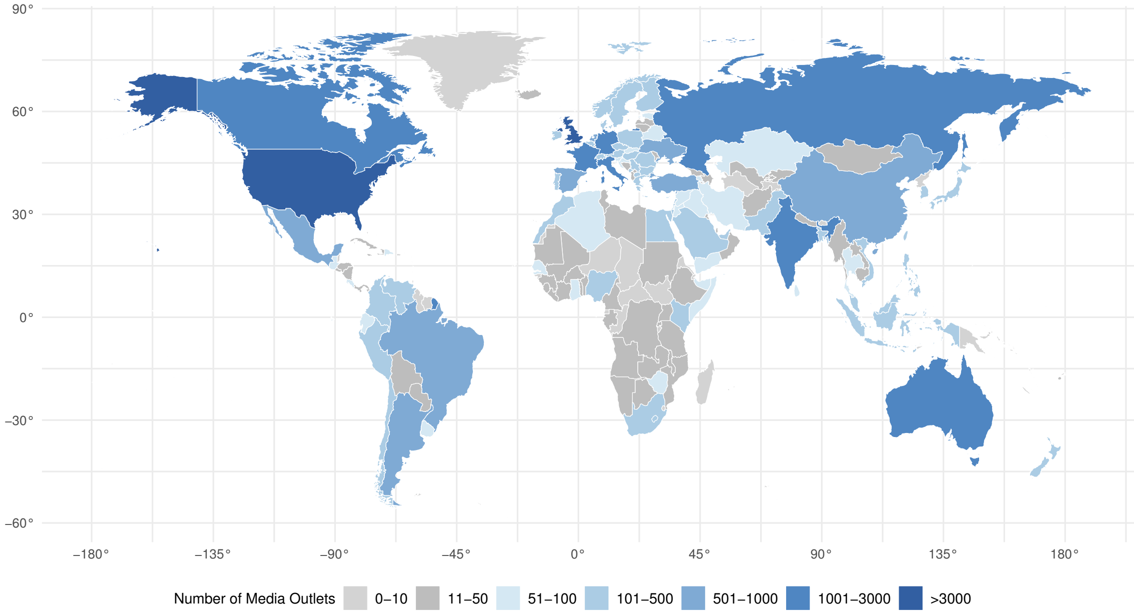
	Dependent Variable: Monthly Average Tone	
	(1)	(2)
Institutional Distance	-0.231***	-0.228***
	(0.043)	(0.041)
Institutional Distance × State-Owned Media Outlet	-0.295***	-0.293***
	(0.079)	(0.079)
Observations	81,288,434	81,288,434
Number of Country Pairs	14,379	14,379
Target Country × Theme × Year FEs	Yes	Yes
Outlet × Year FEs	Yes	No
Outlet × Theme × Year FEs	No	Yes

Notes: This table presents the heterogeneous association between institutional distance and media sentiment by predicted state ownership of a given media outlet. Ownership of each media outlet is labeled by the GPT-4o model. Other variable construction and model specifications follow the convention in Table 1. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

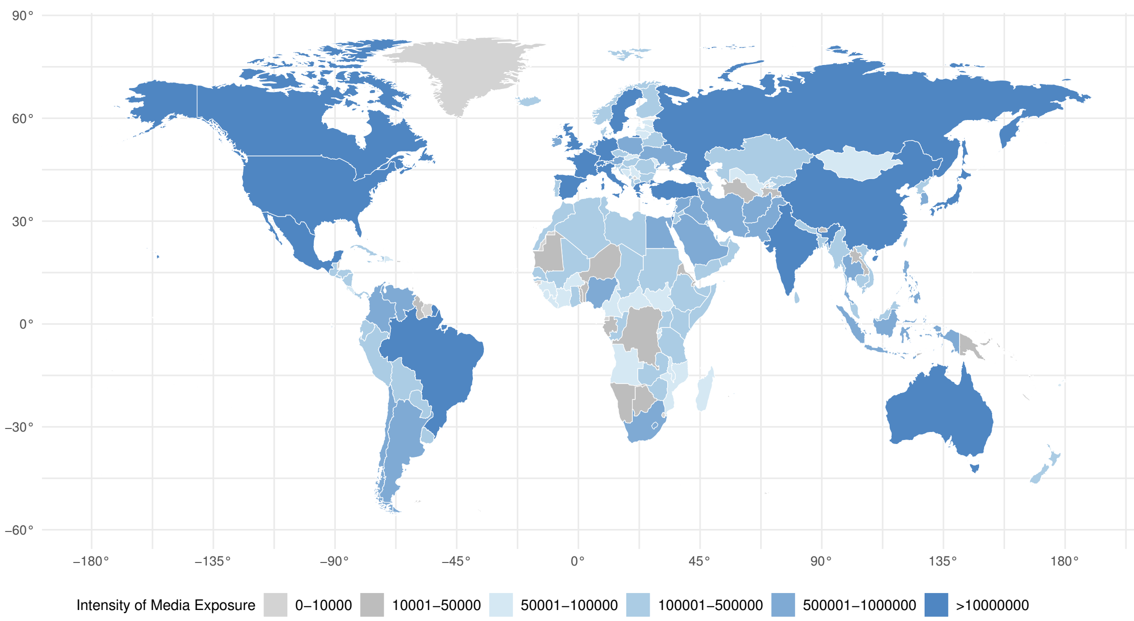
Table A21: Media Alliance and Diplomatic and Economic Alignment

	Dependent Variable		
	UN Voting Distance	Bilateral Investment Treaty	Regional Trade Agreement
	(1)	(2)	(3)
Media Alliance (t-1)	-0.917***	0.219***	0.536***
	(0.054)	(0.030)	(0.035)
Observations	40,954	50,860	50,860
Number of Country Pairs	7,358	8,186	8,186
Source Country FEs	Yes	Yes	Yes
Target Country FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes

Notes: This table reports regressions relating our media alliance measure to subsequent bilateral alignment outcomes. The independent variable is the media alliance index between country i and country j in year $t-1$, constructed from the correlation in their media narratives about third-party countries. The dependent variables are UN General Assembly voting alignment in column (1), the presence of a bilateral investment treaty in column (2), and joint participation in a regional trade agreement in column (3). Details on variable construction are provided in Online Appendix F. Standard errors in parentheses are clustered at the country-pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.



(a) Global Distribution of Media Outlets



(b) Media Exposure

Figure A1: Media Landscape

Notes: This figure illustrates the geographical distribution of media outlets and coverage. Panel A displays the number of media outlets per country in the baseline sample. Panel B depicts the intensity of media exposure across countries.

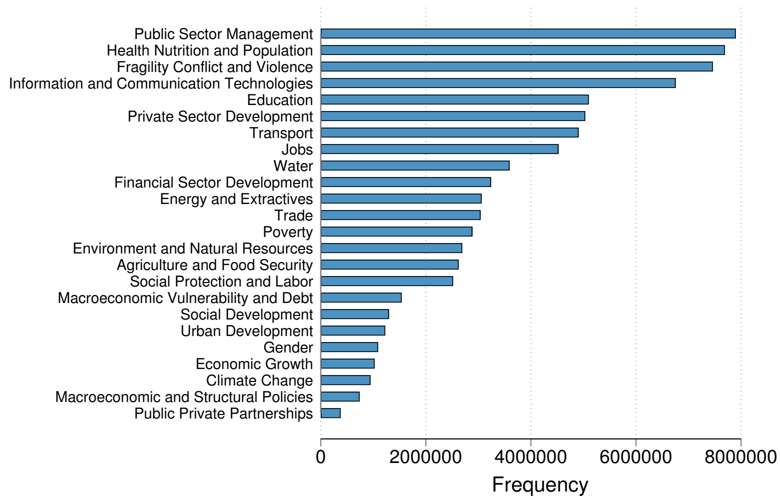
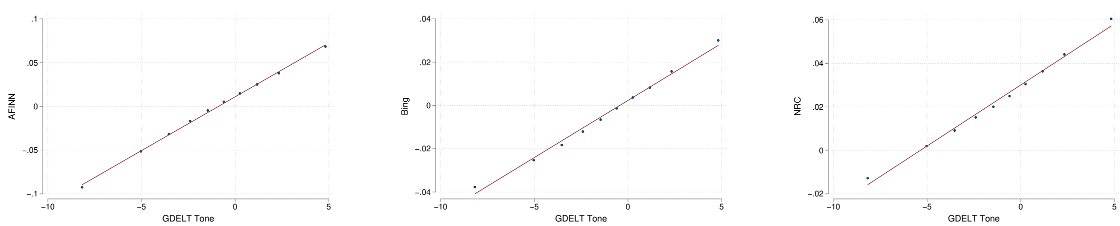


Figure A2: Distribution of Article Themes

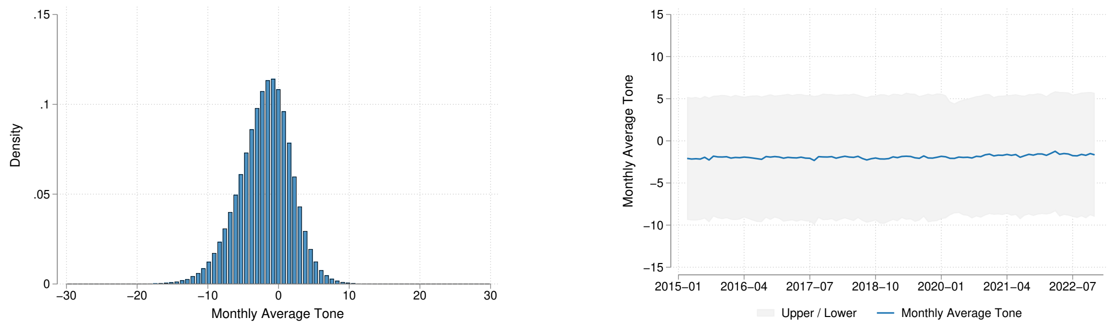
Notes: This figure displays the frequency of observations by theme in the baseline sample. Themes are classified according to Level 1 of the World Bank Group's Topical Taxonomy. If an article addresses multiple themes, it is counted once for each relevant theme.



(a) AFINN (b) Bing Liu et al. (c) NRC

Figure A3: Validation of the Tone Measure

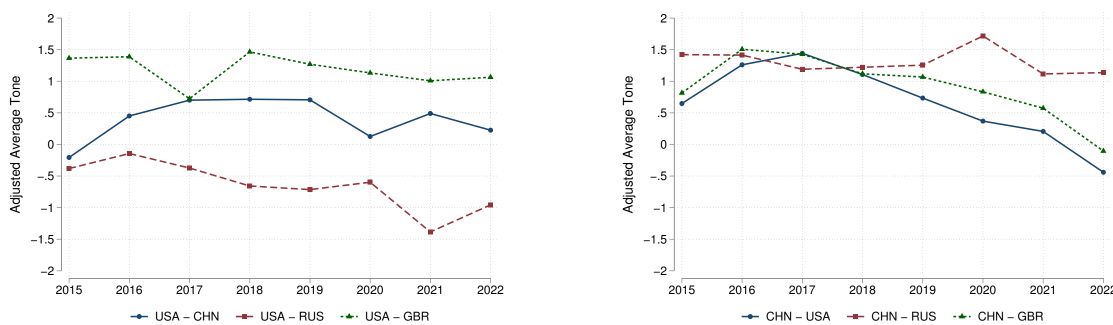
Notes: This figure shows the validation of the GDELT tone variable on a random sample comprising 157,779 news article texts. We construct three sentiment variables based on commonly used lexicon dictionaries, and we plot the correlation between each of these three variables against the GDELT baseline tone measure. We show the scatter plot in ten bins. Panels A, B, and C show the sentiment variable based on the AFINN lexicon, the lexicon from Bing Liu and collaborators, and the NRC lexicon, respectively. The Pearson correlation coefficients range between 0.66 and 0.81.



(a) Tone Distribution (b) Intertemporal Pattern of the Tone Measure

Figure A4: Distribution of Tone Measures

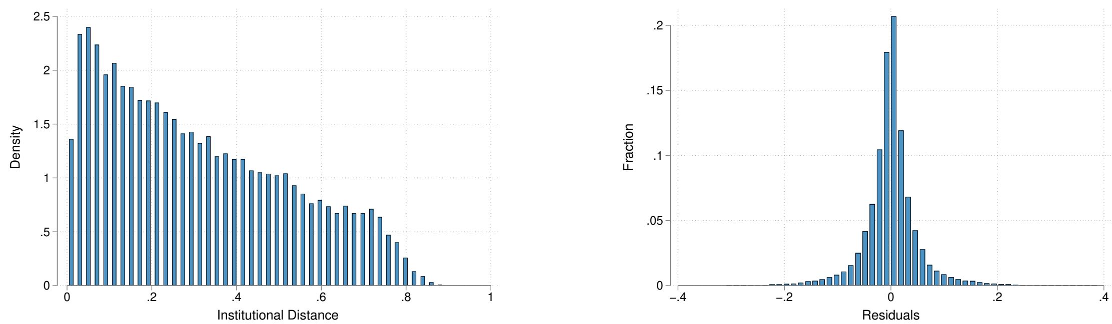
Notes: Panel A presents the distribution of the dependent variable—the monthly average tone at the media-outlet–target-country–theme–month level. Panel B illustrates the temporal pattern of the tone measure, with the 95% confidence interval shaded in gray.



(a) Source Country: USA (b) Source Country: CHN

Figure A5: Tone Measures Over Time by Country

Notes: Panel A displays the average tone of articles published by media outlets in the United States covering issues in China, Russia, and the United Kingdom. Panel B presents the average tone of articles published by media outlets in China covering issues in the United States, Russia, and the United Kingdom. In both panels, tone is adjusted by removing theme-by-year-month fixed effects estimated over the full dataset though results are similar without such adjustment.



(a) Distribution of Institutional Distance (b) Variation Across Years

Figure A6: Distribution of Institutional Distance

Notes: This figure displays the distribution of the main independent variable, institutional distance. Panel A displays the distribution of the raw variable at the country-pair–year level. Panel B displays the distribution of institutional distance residualized by country-pair fixed effects.

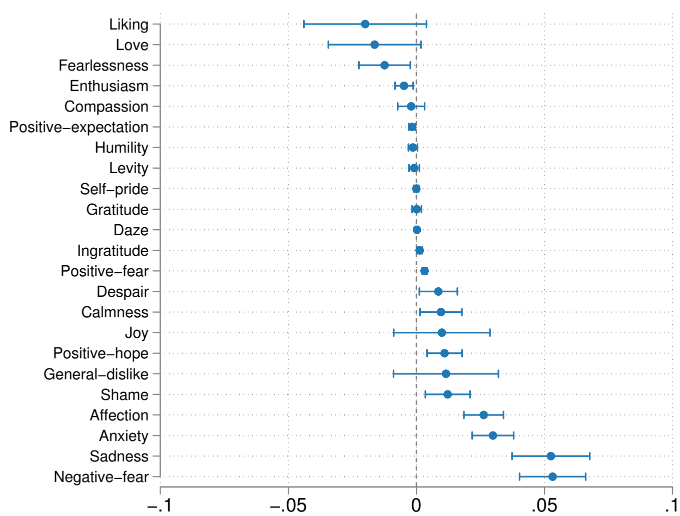


Figure A7: WNA Sentiment Measures and Institutional Distance

Notes: This figure presents the estimated relationship between institutional distance and each sentiment measure individually, as classified by the WordNet-Affect lexicon. The coefficients and 95% confidence intervals are displayed.

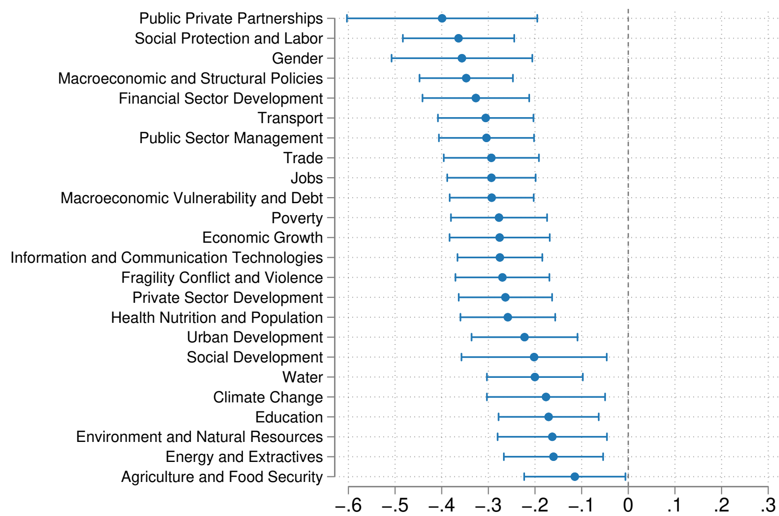


Figure A8: Baseline Estimates by Media Theme

Notes: This figure shows the baseline association between institutional distance and media tone by theme, based on Level 1 of the World Bank's theme classification. The coefficient estimates and 95% confidence intervals are displayed.

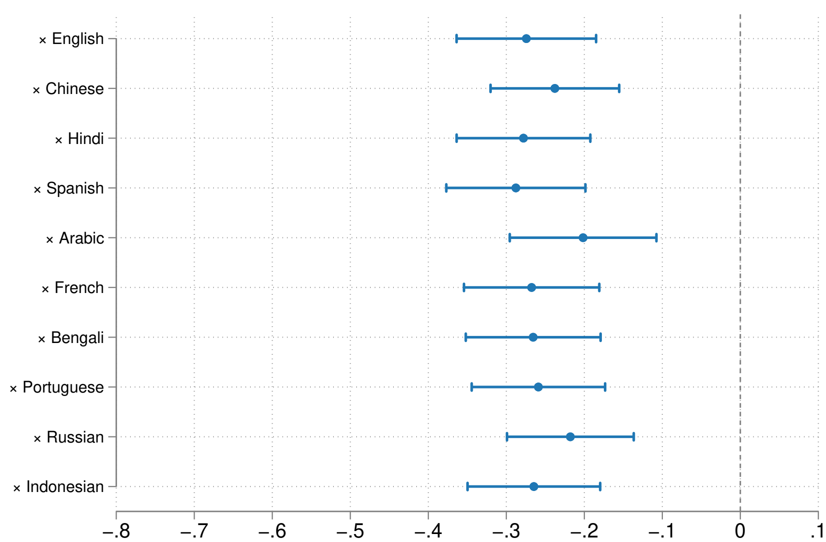
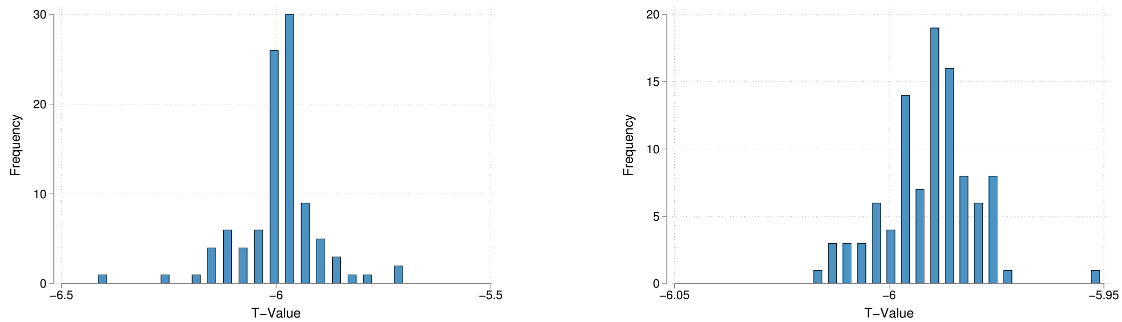


Figure A9: Dropping One Major Language at a Time

Notes: This figure tests the robustness of baseline estimates to excluding one major language at a time, ensuring that the findings are not driven by any single language. The coefficient estimates and 95% confidence intervals are displayed.



(a) Excluding Top 100 Country-Pairs (b) Randomly Excluding 5% Sample

Figure A10: Excluding Subsamples

Notes: This figure assesses the robustness of the baseline estimates to excluding subsets of the sample. Panel A reports t-values for the main coefficient when each of the top 100 country pairs (by number of observations) is excluded in turn. Panel B reports t-values from 100 repetitions in which 5% of observations are randomly dropped.

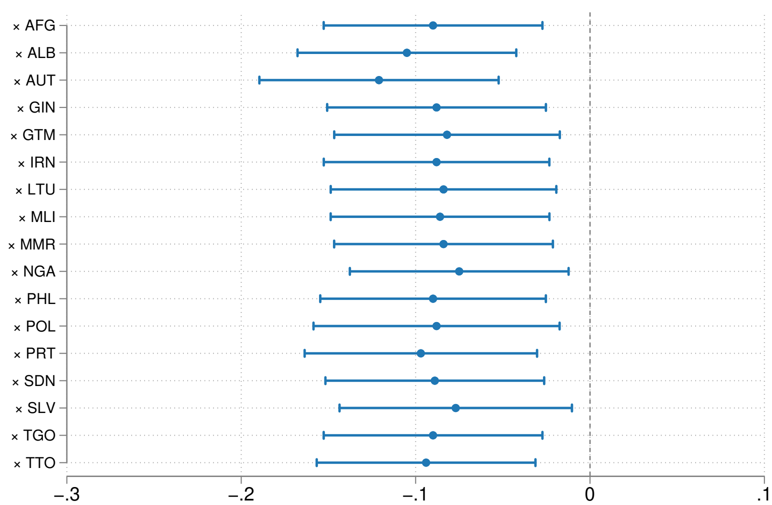
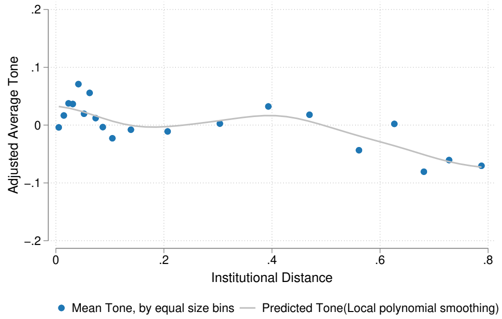
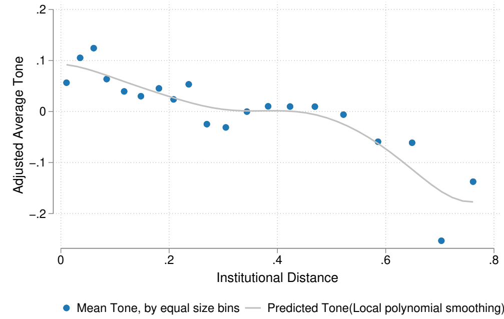


Figure A11: DiD Estimation Excluding One Erosion Episode at a Time

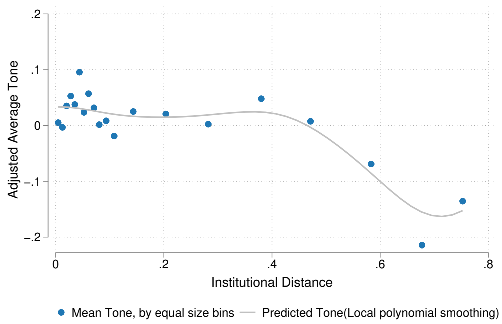
Notes: This figure assess the robustness of the nondymanic DID results to excluding one democratic erosion episode at a time. The coefficient estimates and 95% confidence intervals are displayed.



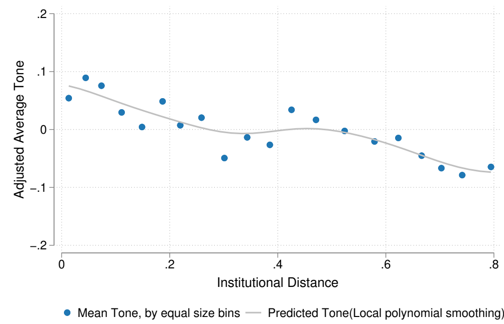
(a) Source Country Is a Liberal Democracy



(b) Source Country Is Not a Liberal Democracy



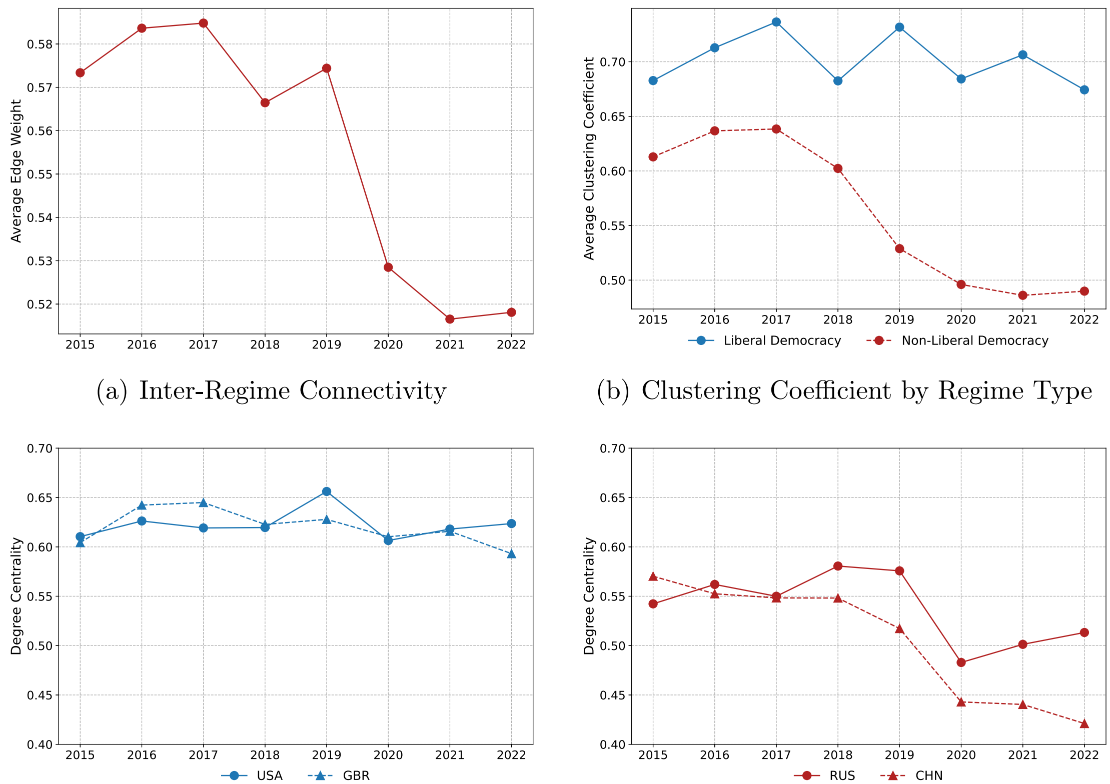
(c) Target Country Is a Liberal Democracy



(d) Target Country Is Not a Liberal Democracy

Figure A12: Democracy and Media Bias, Nonparametric Association

Notes: This figure visualizes the conditional relationship between institutional distance and media bias by regime type. The tone measure is adjusted by removing the two sets of fixed effects used in our baseline analysis—Outlet × Theme × Year, Target Country × Theme × Year—and retaining the residual. The sample is split by Liberal Democracy Indicator$_{c,t}^{\text{Source Country}}$ and Liberal Democracy Indicator$_{c,t}^{\text{Target Country}}$. We group the observations into 20 bins and plot the mean of the adjusted tone against the institutional distance for four separate subsamples.



(a) Inter-Regime Connectivity

(b) Clustering Coefficient by Regime Type

(c) Degree Centrality: USA and GBR

(d) Degree Centrality: CHN and RUS

Figure A13: Structural Changes in Alliance Networks, 2015–2022

Notes: This figure illustrates structural changes in the alliance network from 2015 to 2022. Panel A shows trends in average edge weight between subnetworks of liberal and nonliberal democracies, capturing changes in inter-regime connectivity. Panel B depicts the average clustering coefficient within each group, reflecting the degree of internal cohesion among liberal and nonliberal democracies. Panels C and D present changes in degree centrality for four key countries—the United States, the United Kingdom, China, and Russia—within their respective networks of the same regime type.

B Data Construction

B.1 Construction of the Baseline Sample

B.1.1 Introduction to GDELT GKG 2.0

In contrast to previous research that used data from the GDELT Event 2.0 dataset, we employ data from the GDELT Global Knowledge Graph (GKG) 2.0 for our study. Compared to the GDELT Event 2.0 dataset, the GKG 2.0 offers several benefits that should be highlighted here. First, the Event 2.0 dataset includes only news articles where actors and actions are clearly identified. As of 2022, the Mentions table in the GDELT Event 2.0 dataset contained around 667 million events extracted from 420 million news articles. In contrast, the GKG 2.0 has collected more than 1.4 billion news articles, providing much more detailed information. Second, the GKG 2.0 identifies every country mentioned in a news article, enabling us to focus on news related to a unique target country. Third, as of 2022, the GKG 2.0 has identified over 60,000 themes in news articles using various classification taxonomies. This is more detailed than the relatively ambiguous event types provided by the CAMEO classification used in the GDELT 2.0 Event dataset.

We also replicated our baseline results using data constructed from the the GDELT Event 2.0 dataset. Results displayed in Table A14 show that our baseline results are robust when using a different sample and employing different theme-classification methods based on the GDELT Event 2.0 dataset.

B.1.2 Data Collection and Cleaning

We obtained the GKG 2.0 raw data with over 1.4 billion global news articles from the official GDELT website. We validated and confirmed the data size by comparing it with the summarized result from the GDELT dataset stored in Google Big-Query.

We performed the following data-cleaning steps. First, we restricted the news articles to online articles with distinct URLs to ensure comparability. We removed a small amount of data from broadcasts and journals from our sample. Then, we removed news articles that mentioned more than one country with the V2Locations variable provided in the GKG 2.0 dataset. We did not consider news articles with multiple target countries because the overall tone of a news article would be influenced by multiple targets simultaneously, making it difficult to identify the impact of institutional distance. With these two steps, we retained news articles that reported on a single country and had a unique URL. Around 40% of news articles were left for further analysis.

Next, we aggregated the sample to the media-outlet–target-country–theme–year-month level. We identified the themes covered in each news article using the World Bank Group Topical Taxonomy, which we will discuss in the next section. During the aggregation process, we removed news articles without any theme included in the taxonomy. At the aggregated level, we calculated the monthly average sentiment of news articles from each media outlet that targeted a given country within a specific theme using the V2Tone indicator provided by the GKG 2.0 dataset. This variable records the sentiment of every news article using a dictionary method and has been available since the beginning of the GKG 2.0 project. In contrast, other sentiment scores were included in the GKG 2.0 data in subsequent waves and are not provided for every news article. We validated the reliability of the V2Tone indicator using three widely used sentiment measures, as shown in Figure A3. We also obtain consistent results using the other two sentiment measures in the GKG 2.0 dataset in Table A7.

The aggregated data include around 100 million observations at the media-outlet–target-country–theme–month level. Each observation represents the sentiment from a media outlet toward a target country on a given theme in a specific month. We confirmed the source country of every media outlet (see details in Section B.1.4 below) and dropped media outlets without a clear source country. Then, we removed observations where the media outlet reported on domestic issues. We then further restricted the sample to the eventual baseline of 81,288,434 observations by removing countries without V-Dem data and dropping singletons.

The final baseline sample contains news articles from 55,240 media outlets across 176 source countries reporting on issues related to 177 target countries between 2015 and 2022. 14,379 country pairs are included. Figure A1 displays the global distribution of the source countries and target countries in our baseline sample.

B.1.3 Themes Classification

We identified the themes covered in each news article using the World Bank Group Topical Taxonomy, which consists of around 2,000 entries in the version used by GDELT project. The GDELT project has incorporated the entire World Bank Group Topical Taxonomy into the GDELT GKG 2.0 dataset since March 2015, providing a comprehensive classification typology for news articles. Using this typology, we determined whether any of the themes were mentioned in each news article. It's worth noting that a news article can cover multiple themes simultaneously.

The GDELT project has identified over 60,000 themes from different classification systems, such as the World Bank Group Topical Taxonomy, UN My World Survey Topics, and CrisisLex Taxonomies, in the GKG 2.0 dataset. Among these, we found that the World Bank Group Topical Taxonomy was the most suitable for our research, for two main reasons.

First, this taxonomy was incorporated into the dataset within just one month after the beginning of GKG 2.0 project, whereas other classification schemes were included later. For example, the UN My World Survey Topics were updated in January 2016. Second, the World Bank Group Topical Taxonomy is organized by World Bank teams to record their comprehensive (and even any trivial) work on a global scale. Therefore, it covers a wide range of subjects, from socioeconomic problems to natural issues. In contrast, other classification schemes used by the GKG 2.0 dataset were less systematic or just focused on specific media outlets. For instance, the CrisisLex Taxonomies primarily measure emergent issues.

The World Bank Group Topical Taxonomy comprises 10 hierarchical levels, with 24 Level 1 themes. We obtained the topical structure from the World Bank's website and assigned each theme identified by GKG 2.0 to one or more Level 1 themes. Approximately 66% of news articles from GKG 2.0 dataset contained at least one theme from the World Bank Group Topical Taxonomy.

B.1.4 Identifying the Country of Origin of Media Outlets

A critical empirical challenge we faced was identifying the source countries for tens of thousands of news outlets. This task was complicated by the difficulty of accessing consistent, reliable location data, as well as registration records for media outlets. Even for well-known news outlets, pinpointing their actual location was still challenging, as many outlets operate abroad to circumvent censorship or for other strategic reasons in today's interconnected media landscape. In addition, conventional methods such as IP geolocation through Domain Name System (DNS) records are inadequate, since many outlets outsource server operations to third-party hosting companies or rely on cloud-based hosting services.

To overcome these challenges, we adopted a multistep approach to identify media outlets' source countries. First, we obtained a comprehensive list of news agencies by country from Wikipedia (https://en.wikipedia.org/wiki/List_of_news_agencies). This step ensured that all major media outlets were accurately classified by their country of origin. The list included 187 outlets with corresponding websites, organized by country. Second, we used domain suffix analysis (TLDs) to determine the source country. However, because many media outlets frequently opt for generic suffixes (.com, .net, .org) rather than country-specific domains, this method alone was insufficient for identifying the source countries.

For outlets not identified through the Wikipedia list or TLD analysis, we turned to ChatGPT-4o for country-of-origin identification. It is worth mentioning that the GDELT Project team recommends the "Most Frequency Rule" for similar cases (https://blog.gdeltproject.org/mapping-the-media-a-geographic-lookup-of-gdelts-sources/). This approach infers the source location based on the country with the highest frequency of coverage, under the assumption that news outlets primarily focus on domestic issues. While useful, this method has its limitations. For instance, according to GKG 2.0, the British newspaper *Daily Mail* (www.dailymail.co.uk) covers the United States more often than domestic UK issues.

To ensure greater accuracy, we relied on the ChatGPT-4o API as our primary tool for identifying media outlets' source countries. To validate its performance, we randomly selected 1,000 news outlets from our dataset, manually coded their country of origin, and compared the results with ChatGPT's classifications. The ChatGPT-4o model achieved an impressive 94% accuracy rate in this validation exercise. Further, as a robustness check, we used the most frequently reported country for each outlet as an alternative source identification measure.

In Table A10, we show the baseline results using different source-country identification methods. The estimated effects are comparable across methods.

B.2 Construction of the National Leader News Sample

B.2.1 List of National Leaders

We constructed a comprehensive dataset of de facto supreme leaders across all sovereign states and autonomous regions spanning 2015 through 2022. For initial data collection, the Herre Database on Political Leaders was our primary source (Herre, 2023), supplemented by the Wikipedia list of heads of state and government (https://en.wikipedia.org/wiki/List_of_current_heads_of_state_and_government). To

Top National Leaders by Reporting Frequency

Rank	National Leaders	Country
1	Donald Trump	United States of America
2	Joe Biden	United States of America
3	Boris Johnson	United Kingdom
4	Benjamin Netanyahu	Israel
5	Vladimir Putin	Russia
6	Imran Khan	Pakistan
7	Narendra Modi	India
8	Emmanuel Macron	France
9	Angela Merkel	Germany
10	Jinping Xi	China

Notes: This table presents the top-ten most frequently reported national leaders in our sample.

ensure accuracy, we specifically identified leaders who wielded actual executive authority rather than playing just ceremonial roles, taking into account various political systems, including presidential, parliamentary, and one-party states. For instance, in parliamentary monarchies such as the United Kingdom, we listed the prime minister rather than the monarch, while in one-party states such as China, we included the general secretary of the ruling party.

B.2.2 Identifying Popular Names of National Leaders

We manually coded the various forms of names used by news media outlets to report on national leaders in our dataset. We conducted comprehensive searches on Google News for articles featuring each leader, and we compiled their names (and nicknames) most frequently used by news media. Our focus was on documenting both standardized naming conventions and region-specific variations, including formal titles, common abbreviations, and culturally specific designations.

Using the resulting list, we then identified relevant English-language media coverage in the GDELT dataset through a two-step verification process. Articles were considered relevant to a leader if (i) the article headline contained at least one documented name reference and (ii) the target country of the news article, as identified by GDELT, was the same as the country where the leader holds power. This approach ensured the accurate identification of relevant news coverage.

B.3 Construction of Event Sample

As a parallel project of the GDELT, Event 2.0 extracts events from global news happening in different societies. Like GKG 2.0, Event 2.0 collects information from news articles with diverse languages using machine-learning algorithms on a 15-minute basis. Accordingly, it provides variables for each event, including actors, event types, event locations, and news sentiment. Detailed information on Event 2.0 data is available in the GDELT codebook (https://blog.gdeltproject.org/gdelt-2-0-our-global-world-in-realtime/).

While the event data has limitations—theme classification and restricted data size, as we discussed in Section B.1.1—it remains valuable for robustness checks. Following the data-cleaning procedures we applied to the GKG 2.0 Dataset, we generated country-pair reporting data from the Event 2.0 dataset through the following steps.

First, we identified the target country of each event using location variables. The Event 2.0 dataset provides location information for both actors involved in each event. We extracted the unique address values from these location attributes to designate the target country. Consequently, events with multiple locations and target countries are excluded from the final analysis.

Next, we mapped the source country for each media outlet using the same method employed in our main results. Finally, we aggregated the data at the media-outlet–target-country–event-type–month level and calculated the average tone of the news articles. Event 2.0 classifies news themes using a variable called *event type,* based on the coding scheme from the Conflict and Mediation Event Observations (CAMEO) framework. We applied the top level of this classification scheme, which includes 20 event types. We utilized these aggregated data for the robustness check; the results are presented in Table A14.

B.4 Sample Construction for Public Opinion Responses: Gallup World Poll

We collected public opinion data from the Gallup World Poll, one of the most influential international surveys. Starting in 2006, Gallup has conducted ongoing surveys of people in over 160 countries and areas, representing more than 98% of the global adult population. To do this, they randomly select nationally representative samples covering around 3 million residents. Generally, Gallup surveys 1,000 people in each country or region for each wave, employing a consistent set of core questions translated into the primary languages of each respective country.

The poll aims to identify the most urgent challenges confronting humanity and to monitor advancements in human development, including questions related to business and economics, government and politics, and public health. These indicators help researchers understand the broad context of national interests and establish specific relationships between indexes and lagging economic outcomes.

The Gallup World Poll asks respondents to evaluate their approval of foreign leaders using a binary agree/disagree response. We exclude respondents who refuse to answer or report "don't know." This variable helps us explore the relationship between cross-country media bias and the sentiments of local people. To avoid the reverse-causality problem, we used the approval variable one year after the media reports for each observation. In the analysis, we also include gender, income level, educational level, and employment status as control variables.

The 2016–2023 poll we used in our sample surveyed people's approval of leadership in nine target countries: China, Germany, France, India, Iran, Japan, Russia, Turkey, and the United States. The respondents are from 147 countries and areas. In total, 779 country pairs are identified. Summary statistics of this sample, in the table below, support the results in Table 7.

	N	Mean / %	Std	Median
Individual level sample				
Approval of Leadership	3,570,351	0.521	0.500	1
Gender (1 = Female)	3,570,351	50.44%	-	-
Income Level	3,570,351	3.268	1.415	3
Education Level	3,570,351	1.957	0.680	2
Unemployment (1 = Unemployed)	3,570,351	6.79%	-	-
Country-pair-Year level sample				
Tone	4,665	-1.421	1.263	-1.435

C Examples of News Articles From GDELT

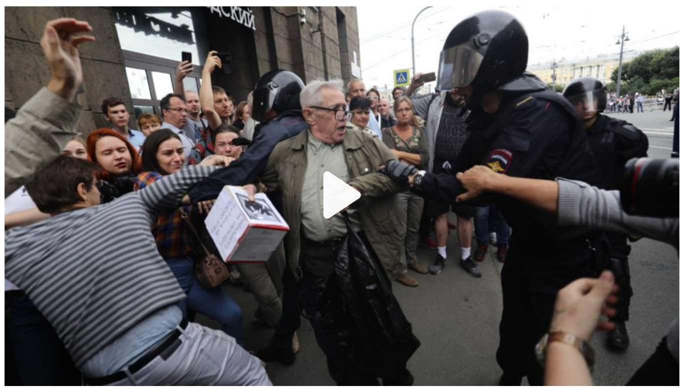
Example 1

Below is a news article from CNN collected by the GDELT project that focuses on Russian politics. The tone of this news article is -3.88.

Vladimir Putin is not invincible

Analysis by Nathan Hodge, CNN

3 minute read · Updated 5:53 AM EDT, Thu September 27, 2018



Example 2

Below is a news article from China.org.cn collected by the GDELT project that highlights the problem of gun violence in the United States. The tone is - 7.78.

Research on gun violence thwarted in U.S.: report

Xinhua, June 25, 2022 0 Comment(s) Print E-mail

Adjust font size:

NEW YORK, June 24 (Xinhua) -- For more than 20 years, the U.S. federal government failed to properly support research into gun violence, reported Scientific American on Thursday.

"We spent about 63 U.S. dollars in research dollars per life lost to gun violence, compared to roughly 1,000 dollars per life lost to car crashes and nearly 7,000 dollars per life researching sepsis, a life-threatening response to infection," said the report.

"Two decades without these critical investments has left us myriad open questions about the effects of gun policies," it noted.

More than 45,000 people died from firearm injuries in 2020. Firearms have also become the leading cause of death for children, according to the report.

"These deaths come amid a surge in the domestic sale of firearms. The most recent Small Arms Survey estimates there are nearly 400 million guns in the United States," it added. Enditem

*Follow China.org.cn on **Twitter** and **Facebook** to join the conversation.*

ChinaNews App Download

D Major News Producers by Country

	United States	United Kingdom	China	Russia
1	yahoo.com	reuters.com	xinhuanet.com	sputniknews.com
2	msn.com	dailymail.co.uk	china.org.cn	regnum.ru
3	reliefweb.int	bbc.com	sina.com.cn	rt.com
4	humanitariannews.org	theguardian.com	eastmoney.com	rg.ru
5	prnewswire.com	lexology.com	cnfol.com	ria.ru
6	sfgate.com	bbc.co.uk	people.com.cn	tass.ru
7	marketwatch.com	trust.org	peopledaily.com.cn	lenta.ru
8	chron.com	mondaq.com	sina.com	runews24.ru
9	businessinsider.com	mirror.co.uk	china.com	riafan.ru
10	washingtonpost.com	aawsat.com	cfi.net.cn	life.ru

Notes: This table presents the top 10 most frequent news sources from four major countries. The ranking is based on the total number of observations in our baseline sample. Because GDELT indexes machine-discoverable URLs, high-volume aggregators and wire services are overrepresented, while major national newspapers (such as the *New York Times*) may appear less frequently due to lower publication volume.

E Robustness Checks and Additional Results

E.1 Additional Results

Estimates by year. To track the evolution of media bias over time, we first augment our baseline model by interacting institutional distance with a linear time trend. As shown in columns (1) and (2) of Table A5, the degree of media bias increases over time. The coefficient grows annually by 12% (=0.023/0.191 as in column (2)) relative to the baseline level. Next, we estimated a model allowing the coefficient on institutional distance to vary flexibly by year:

$$\text{Tone}_{i(c),j,s,m(t)} = \beta_t \,\text{Institutional Distance}_{c,j,\,t-1} + \delta_{i(c),s,t} + \theta_{j,s,t} + \varepsilon_{i(c),j,s,m(t)}. \quad \text{(A1)}$$

Other terms follow the definitions in Equation (4). The results are shown in columns (3) and (4) of Table A5. Consistent with the trend-interaction results, the association between institutional distance and tone becomes steadily more negative over time. These patterns underscore the growing relevance of political-system differences in shaping global media bias and suggest that institutionally distant countries are becoming increasingly more negative toward one another in their news coverage.

Estimates by theme. A key question is whether media bias is confined solely to political topics, which are commonly believed to be the most sensitive to institutional differences. To demonstrate the effect of institutional distance by media theme, we estimated the following model allowing the main coefficient to vary by theme:

$$\text{Tone}_{i(c),j,s,m(t)} = \beta_s \,\text{Institutional Distance}_{c,j,\,t-1} + \delta_{i(c),s,t} + \theta_{j,s,t} + \varepsilon_{i(c),j,s,m(t)}. \quad \text{(A2)}$$

Other model specifications follow the definitions in Equation (4). We plot the estimated theme-specific coefficients in Figure A8. Indeed, themes directly linked to the government, such as "Public-Private Partnerships" and "Public Sector Management," exhibit some of the strongest bias by institutional distance. Yet media bias extends well beyond public-sector-related themes, affecting a wide range of socioeconomic issues, such as "Social Protection and Labor," "Gender," "Financial Sector Development," "Economic Growth," and "Trade."

E.2 Robustness

We conduct an extensive battery of robustness checks for the baseline Table 1.

First, we examine whether our results are sensitive to alternative model specifications. In columns (1) and (2) of Table A6, we introduce country-pair fixed effects absorbing all time-invariant confounders at the country-pair level, including any time-invariant distance measures. Since institutional distance generally evolves rather slowly, this specification absorbs part of its variation, leading to a somewhat smaller coefficient. Nevertheless, the estimates remain robust and statistically significant. In columns (3)–(5), we make the baseline fixed effects even more granular by including outlet × theme × year-month and target country × theme × year-month fixed effects. The results remain stable under these demanding specifications.

Second, we considered alternative measures of both the dependent and independent variables used in our baseline analysis. Table A7 presents results using two alternative tone measures based on the Lexicoder Sentiment Dictionary (LSD; Young and Soroka, 2012) and SentiWordNet 3.0 (SWN; Baccianella, Esuli, and Sebastiani, 2010) provided in the GKG 2.0. The results are consistent with our baseline findings. Table A8 further demonstrates that the findings are also robust to two alternative measures of institutional distance, constructed using the V-Dem Electoral Democracy Index and the Polity Index. Importantly, the estimated magnitudes are very similar across these distance measures when expressed in standard-deviation units.

Another measurement-related concern is that sentiment measures derived from GDELT's machine-translated text may be noisy, and bias could be associated with the quality of the translation. We randomly sampled 50,000 articles per month from the GDELT GKG 2.0 database to construct an article-level dataset, and we show that a translation dummy variable (equal to 1 if the article is machine-translated) is not significantly correlated with article-level tone. In our baseline sample, we also constructed a variable capturing the proportion of English-language articles, and we find that the estimated coefficient on institutional distance remains stable when controlling for the proportion of English-language articles. Furthermore, the interaction between the share of English-language articles and institutional distance is not statistically significant. We further restricted the sample to countries where one of the world's ten most widely spoken languages is official, addressing concerns that the results may be driven by translation quality in low-resource language settings. The results remain fairly robust. These results are reported in Table A9. We also iteratively excluded each major language from the sample, and we find that the results remain largely unaffected (Figure A9).

Third, we employed a range of alternative rules to identify each outlet's source country. Table A10 shows that our estimates are stable across all such complementary approaches—whether we assign countries by the Most Frequency Rule (the country the outlet covers most often), DNS lookups, ChatGPT classifications, or based on top-level domain suffixes (e.g., ".uk" for the United Kingdom). Limiting the sample to 187 major outlets whose home countries are explicitly identified by Wikipedia more than doubles the estimated coefficient, even though the sample size falls sharply. Focusing on these prominent, widely recognized outlets thus reveals an even stronger media bias toward institutionally distant countries, further reinforcing our baseline findings.

Fourth, we conducted additional tests using alternative samples. We report the results based on the tone variable winsorized at 0.5%, 1%, and 5%, respectively, in Table A11. Given that news articles published by U.S. media outlets or covering U.S.-related issues constitute a significant portion of our sample, we performed a robustness check by excluding all observations related to the United States. The results, reported in Table A12, are quantitatively similar to our baseline estimates. In addition, Figure A10 suggests that our findings are unlikely to be driven by a small fraction of the sample. Panel A displays the t-values for the coefficient of interest when excluding each of the top 100 country-pairs ranked by number of observations. Panel B presents results based on randomly excluding 5% of the observations in each iteration. Further, we aggregated the baseline sample to the country-pair–theme–year level, and the results remain consistent with our main findings (Table A13).

Fifth, we replicated our baseline results on an alternative sample based on the GDELT Event 2.0 dataset. In Online Appendix B, we discuss at length why we preferred the GKG 2.0 to the Event 2.0 database. The results obtained from the GDELT Event 2.0 sample are reported in Table A14. Unlike the baseline specifications, this analysis incorporates event-type fixed effects instead of theme fixed effects, as the GDELT Event 2.0 dataset does not include theme classifications for news articles. Nevertheless, the statistical significance and magnitude of the estimated coefficients remain consistent with our baseline findings. These results suggest that our findings are robust to different sample-construction and theme-classification methods.

We further controlled for economic interdependence and conflicts between the source and target countries. Specifically, in Table A15, we control for the source country's trade dependence on the target country (column 1), its trade balance with the target country (column 2), and a binary indicator for whether the source country was subject to sanctions imposed by the target country (column 3). While the inclusion of these variables may risk overcontrolling—since we interpret institutional distance as partly reflecting underlying geopolitical divergence between countries—we find that none of them are statistically significant, and the institutional distance coefficient remains robust and largely unaffected.

Finally, our findings are robust to alternative standard-error clustering methods, as shown in Table A16. In the baseline estimation, we clustered the standard errors at the country-pair level. We also show that coefficients are consistently significant when clustering at the media-outlet or source-country level. Following Petersen (2008), we also employed two-way clustering methods. We obtain similar results when standard errors are clustered at the (pair, year), (outlet, year), or (source country, year) levels. Reassuringly, the results hold up under any of the clustering methods considered.

F Variable Definitions

Panel A: Variables at the Media-Outlet–Target-Country–Theme–Month Level (Baseline Sample)

Variable	Definition	Source
Monthly Average Tone	Monthly average tone of news article. Tone of each news article is calculated with the difference between positive tone and negative tone.	GDELT GKG 2.0
Monthly Average Positive Tone	Monthly average positive tone of news articles. Positive tone of each news article is measured with the percentage of all positive words used in the article.	GDELT GKG 2.0
Monthly Average Negative Tone	Monthly average negative tone of news articles. Negative tone of each news article is measured with the percentage of all negative words used in the article.	GDELT GKG 2.0
Monthly Average Tone (LSD)	Monthly average tone of news article measured with the Lexicoder Sentiment Dictionary (LSD). Tone (LSD) of each news article is calculated with the difference between positive tone (LSD) and negative tone (LSD). Positive/Negative tone (LSD) is measured with the percentage of positive/negative words used in a article.	GDELT GKG 2.0
Monthly Average Tone (SWN)	Monthly average tone of news article measured with the SentiWordNet 3.0 (SWN) scores. Tone (SWN) of each news article is calculated with the difference between positive tone (SWN) and negative tone (SWN) .	GDELT GKG 2.0

Panel B: Variables at the Country-Pair–Year Level

Variable	Definition	Source
Institutional Distance	Dissimilarity of political institutions between two countries in the past year, measured by the difference between their scores in the liberal democracy index (v2x_libdem) provided by the V-Dem Project. This index assesses to what extent the ideal of liberal democracy is achieved in a country.	V-Dem Project
Institutional Distance (Electoral Democracy)	Dissimilarity of political institutions between two countries in the past year, measured by the difference between their scores in the electoral democracy index (v2x polyarchy) provided by the V-Dem Project. This index assesses to what extent is the ideal of electoral democracy in its fullest sense achieved in a country.	V-Dem Project
Institutional Distance (Polity2 Score)	Dissimilarity of political institutions between two countries in the past year, measured by the difference between their scores in the Polity 2 index. This index assesses the general democratic level in a country.	Polity V
UN Voting Distance	Dissimilarity of voting preference between two countries in the past year, measured by the difference between their ideal point estimates derived from the United Nations General Assembly (UNGA) Voting Data.	Voeten, Strezhnev, and Bailey (2023)

Panel B: Variables at the Country-Pair–Year Level (cont'd)

Variable	Definition	Source
Common Language	A dummy variable indicating whether two countries share the same official language.	CEPII
Geographic Distance	Geographic distance between the capitals of two countries.	CEPII
Religious Distance	Religious dissimilarity between two countries generated from a principal component analysis using three indicators.	Dow and Karunaratna (2006)
Genetic Distance	Genetic dissimilarity between two countries, measured by the distance to the most recent common ancestors of two populations, or equivalently, the length of time since two populations split apart.	Spolaore and Wacziarg (2009)
Trade Dependence	Trade dependence of country i on country j in the past year, with data from UN Comtrade. It is determined by summing the exports from country i to country j and the imports from country j to country i, then dividing this sum by the total exports and total imports of country i.	UN Comtrade
Trade Balance	The balance of trade for country i with respect to country j in the past year, with data from UN Comtrade. It is determined by subtracting the imports from country j to country i from the exports from country i to country j, then dividing this difference by the GDP of country i.	UN Comtrade
Trade Sanction	A dummy variable indicating whether country i was sanctioned by country j in the past year, with data from the Global Sanctions Data Base (GSDB).	GSDB
Bilateral Investment Treaty	A dummy variable indicating if there is a bilaterater inverstiment treaty between two countries, with data from Electronic Database of Investment Treaties (EDIT).	EDIT
Regional Trade Agreement	A dummy variable indicating if there is a multilateral and bilateral regional trade agreement (RTA) between two countries, with data from Mario Larch's Regional Trade Agreements (RTA) Database.	RTA-Dataset

Panel C: Variables at Country-Year / Year-Month Level

Variable	Definition	Source
Liberal Democracy Index	An index (v2x_libdem) assessing to what extent is the ideal of liberal democracy achieved in a country in the past year.	V-Dem Project
Liberal Democracy Indicator (Yes=1)	A dummy variable indicating whether a country has been a liberal democracy or not in the past year, re-coded from the Regime of the World (Row) measures provided by the V-Dem Project.	V-Dem Project
Democratic Erosion	A dummy variable indicating whether a country experienced a democratic backlash in or before the current year.	V-Dem Project
Freedom of Traditional Media	An index (v2mecenefm) from the V-Dem Project assessing whether the government directly or indirectly attempted to censor the print or broadcast media in the past year. A higher value means that traditional media had higher freedom.	V-Dem Project
Freedom of the Internet	An index (v2mecenefi) from the V-Dem Project assessing to what extent the government attempted to censor information (text, audio, or visuals) on the Internet in the past year. A higher value means that Internet media had higher freedom.	V-Dem Project
State-owned Print Media	An index (v2medstateprint) from the V-Dem Project assessing how many of the top-four national print media with the highest readership in the past year are state-owned.	V-Dem Project
State-owned Broadcast Media	An index (v2medstatebroad) from the V-Dem Project assessing how many of the top-four national broadcast media with the largest audience in the past year are state-owned.	V-Dem Project
Monthly Growth in Nighttime Light Intensity	This measurement is determined by dividing the average nightime light intensity of the past month by the average nighttime light intensity from two months earlier.	VIIRS-DNB
High Inflation Indicator	This indicator is equal to one when the inflation rate over the past year is larger than 0.05. Inflation (annual %) rate in the past year is measured by the "NY.GDP.DEFL.KD.ZG" index from the World Bank.	World Bank
Domestic Turmoil Index	Share of domestic government-targeting events with negative Goldstein sentiment among all domestic government-targeting events in the month, following Amarasinghe (2022) and using data from the GDELT Event 2.0 database.	GDELT Event 2.0

Panel D: Variables at the Media-Outlet Level

Variable	Definition	Source
State Media	A dummy variable indicating the ownership status of the media outlet, whether state-owned or not, labeled by GPT 4o API.	GPT 4o