The War of Ideas: Institutions and Global Media Bias *

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

First Draft: December 31, 2023
This Version: November 30, 2025

Abstract

We present the first systematic analysis of how countries cover each other in the news, using a newly constructed dataset of 55,240 distinct media sources from 176 countries between 2015 and 2022, drawing upon an extensive corpus of 1.4 billion online news articles. Our findings reveal a pattern of global media bias, whereby countries with greater differences in political regimes systematically cover each other in a more negative light, shaping both the sentiment and thematic content of news articles. This bias is particularly pronounced when less democratic countries report on liberal democracies, when state-controlled media is involved, or when the reporting country faces an economic downturn or inflation. Our findings are further corroborated by a natural experiment of democratic erosion episodes among the reporting countries. Additional analysis suggests that cross-country coverage has relevance for equity markets as well as public attitudes toward target countries. Finally, we show that similarities in how countries report on *third* countries can be used to uncover a granular measure of de facto media alliances and ideological blocs.

Keywords: Media Bias, Ideology, Democracy, Geopolitics
JEL Codes: P00, F51, D74, L82

*We are grateful to Daron Acemoglu, Giulia Caprini, Annalí Casanueva Artís, Ruben Durante, Ruben Enikolopov, Germain Gauthier, Pierre-Guillaume Méon, Andrea Prat, Jesse Schreger, Jesse Shapiro, David Yang, and participants at the CEPR Workshop on Media, Technology, Politics, and Society, CESifo Venice Workshop on Digital Platforms, ifo Big Data Seminar, the 2025 APSA Annual Meeting, Hong Kong Polytechnic University, and Lingnan University for insightful comments.

[†]School of Business, Hong Kong Baptist University (email: siboliu@hkbu.edu.hk)
[‡]MIT Sloan School of Management and CEPR (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 differed sharply from one outlet to another. *"Russia invades Ukraine in largest European attack since WWII,"* declared Fox News, conveying a sense of historical gravity and casting Russia as the clear aggressor.[1] Across the Atlantic, The Guardian similarly ran the front page headline *"Putin invades."*[2] Meanwhile, China's state-run media, People's Daily, led with *"Russia-Ukraine conflicts intensify, sending shock waves across world,"* a more neutral headline that avoided explicit language of aggression.[3] On the flip side, The New York Times' coverage of the war in Yemen—waged by the 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 national 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 (Durante and Knight, 2012; Gentzkow and Shapiro, 2010; Groseclose and Milyo, 2005; Larcinese, Puglisi, and Snyder, 2011; 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 remain open questions.

[1] Fox News. February 24, 2022. "Russia Invades Ukraine in Largest European Attack Since WWII." Available at: https://www.foxnews.com/world/russian-invades-ukraine-largest-europe-attack-wwii (last accessed on May 21, 2025).

[2] The Guardian, February 24, 2022. "Putin Invades." Available at: https://pbs.twimg.com/media/FMZJoxsXoAE6Wsk?format=jpg&name=4096x4096 (last accessed on May 21, 2025).

[3] People's Daily Online, February 26, 2022. "Russia-Ukraine Conflicts Intensify, Sending Shock Waves Across World." Available at: http://en.people.cn/n3/2022/0226/c90000-9963162.html (last accessed on May 21, 2025).

This paper presents the first systematic analysis of how countries cover one another in the news, organized around three key research questions: (i) Are there systematic patterns of global media bias in cross-country coverage, particularly along political and ideological lines? (ii) If so, what factors shape the strength and direction of such bias? (iii) Does biased international news coverage have relevance for key economic and political outcomes?

Answering these questions is empirically challenging, for at least three reasons. First, there is a lack of standardized datasets that systematically track how media outlets in different countries report on foreign countries. Building a comprehensive dataset that includes a wide range of countries and media sources is therefore essential for identifying patterns and dynamics of media bias at the international level. Second, media landscapes vary widely across countries in terms of structure, language, and cultural context, making systematic cross-country comparisons difficult. As such, developing a standardized and scalable measure of media bias across countries and outlets is crucial for meaningful empirical analysis. Third, disentangling the key drivers of media bias is a challenging task because media bias can be influenced by a multitude of factors, including but not limited to, the political leanings of individual media outlets, the cultural and political environments of the source country, and the thematic focus of the news coverage.

To address these challenges, we draw on an extensive corpus of 1.4 billion news articles and 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 database of the GDELT Project, which offers real-time translation, as well as emotion and theme detection in 65 languages. Importantly, the GKG 2.0 database provides a tone score for each article, capturing the overall sentiment of its content. Following Puglisi and Snyder (2015), we treat systematic variation in tone as a form of media bias—that is, we consider coverage to be biased when an outlet *systematically* portrays another country in a more favorable or unfavorable light. We aggregate the article-level data to the outlet–target-country–theme–month level, which constitutes the primary unit of analysis in our empirical work.

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 where we regress the average tone of a media outlet toward each foreign country, within a given theme and month, on the *institutional distance* between the two countries in the previous year.[4] We measure institutional distance as the absolute difference in their liberal democracy scores from the Varieties of Democracy (V-Dem) project (Coppedge et al.,

[4]We use the lagged measure to minimize concerns about reverse causality, though our results are similar when using a contemporaneous measure.

2023), which we interpret not only as the difference in regime types but also as a proxy for ideological and political divergence.[5]

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, or global thematic trends. 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 after controlling for other 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, international trade flows (Kleinman, Liu, and Redding, 2024), and country-pair fixed effects that absorb any 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.

We next examine whether 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 idelogical 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.

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 more focused subset of news

[5]Our findings are robust to alternative measures of institutional distance, such as the Polity V democracy index. We also examined media bias by indicators of geopolitical alignment, such as ideal point distance in United Nations General Assembly (UNGA) voting (Bailey, Strezhnev, and Voeten, 2017; Voeten, 2000, 2013). 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.

articles that mention national leaders—a domain widely recognized as politically 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.

We next examine how the relationship between institutional distance and media bias varies with the regime type of the source and target countries. We find that the association is weaker when the source country is a liberal democracy, but significantly stronger when the target country is a liberal democracy. This asymmetry suggests that global media bias is particularly pronounced in how less democratic regimes portray liberal democracies.

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, media bias in cross-national coverage is significantly more pronounced when the source country exhibits higher levels of traditional and internet censorship. Second, it is exacerbated by state ownership of media channels, including both print and broadcast outlets. Third, it intensifies during periods of economic distress in the source country—such as months of economic downturns or surging inflation. Fourth, media bias is stronger during domestic political turmoil in less democratic source countries. These patterns are consistent with an information-control interpretation: governments, particularly authoritarian ones, face stronger incentives to manipulate international narratives during times of domestic stress in order to manage public perception 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 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. We find that, following democratic backsliding in a given country, its media coverage of highly democratic countries becomes significantly more negative relative to coverage of less democratic countries. We find no evidence of pre-trends, 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.

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, while less democratic regimes form separate subclusters, particularly around Russia and China.

Finally, we explore the relevance of global media bias for economic and political outcomes.

First, we analyze its informativeness for financial market behavior. We construct a sample of U.S.-listed firms with detailed data on revenue shares by foreign country and compute firm-month–level exposure to negative media tone toward the United States. We find that firms with greater sales exposure to countries issuing more negative coverage of the U.S. in a given month tend to experience immediate declines in stock prices. The pattern suggests that foreign media sentiment may influence perceptions or serve as useful indicators of political or economic risk, with significant implications for investment decisions.

Second, we examine the relationship between global media bias and public opinion. Using Gallup World Poll data, we find that more favorable sentiment in a country's media coverage of another country is associated with an increase in public favorability toward the latter in the following year. While not necessarily causal, this pattern is consistent with the idea that international media narratives can shape, or at least anticipate, broader shifts in public opinion, highlighting the potential political relevance of cross-border media sentiment.

This paper contributes to the literature in several ways. First, it advances our understanding of the interplay between media and politics (for reviews, see Prat and Strömberg, 2013; Strömberg, 2015; Zhuravskaya, Petrova, and Enikolopov, 2020), with a particular focus on media bias. While 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), our study offers the first systematic analysis of global media bias—that is, how countries' media outlets differentially cover other nations depending on their ideological and political alignment. Some prior work has examined specific cases; for example, Chen and Han (2022) shows that UK- and US-based outlets adopted a more negative tone toward China after being blocked by the Chinese government.[6] In contrast, our analysis provides a comprehensive, comparative perspective

[6]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), Baum and Zhukov (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; and 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.

across a broad range of countries and thematic domains. By shedding light on this largely overlooked international dimension, our paper expands the scope of media bias research and fills an important gap in the literature.

The findings of our paper closely align with a burgeoning literature on information control in non-democratic 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).[7] 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, 2013, 2017; Roberts, 2018) and leverage state ownership of media outlets 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 distress. These patterns underscore the role of international media coverage as a strategic—yet often overlooked—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.[8] Building on this insight, we use patterns in global media coverage to develop 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; Mohr and Trebesch, 2024; Broner et al., 2025). By showing that sentiment in international news coverage is informative

[7] 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 (Besley and Prat, 2006; Durante and Knight, 2012; McMillan and Zoido, 2002; Qian and Yanagizawa-Drott, 2017; Tella and Franceschelli, 2011).

[8] 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).

for key financial outcomes such as stock prices and cross-border lending, we demonstrate that media narratives can serve as timely indicators of shifts in international relations.

The remainder 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 is sourced from the Global Database of Events, Language, and Tone (GDELT) (Leetaru and Schrodt, 2013). The GDELT Project consists of two main databases: the GDELT Event Database and the GDELT Global Knowledge Graph (GKG). The Event Database focuses on recording *individual events* by aggregating data from various news articles and 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 and content, allowing for a detailed analysis of their themes, contexts, and sentiments. This 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, images, and events. A detailed overview of the GKG Database 2.0 is provided in Online Appendix B, along with two illustrative examples of news articles from our database in Online Appendix C.

Figure 1 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 areas indicate regions with fewer outlets. The United States stands out as the country with the highest number of media outlets, whereas many African countries have a more limited presence of media outlets 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 in news reporting.

We further illustrate the bilateral reporting relations of the database in Figure 2. 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 1 and 2 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 multi-step 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 more detailed discussion of this 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 other country-level datasets containing 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, violence, among others. To provide a more granular clas-

sification, each level-1 theme is further divided into sub-themes. In our research, we utilize 24 level-1 themes from the WBG Topical Taxonomy as reported in the GKG 2.0 database. Figure 3 presents the distribution of the themes. The themes most frequently covered in news articles include "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 used 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 official 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. Each news article was categorized into one of the 24 themes defined by the World Bank Group Topical Taxonomy. To focus specifically on cross-border reporting, we excluded observations covering domestic news.

In our final sample (outlet-by-target-country-by-theme-by-month), 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 the "tone" and "sentiment" in news coverage (Puglisi and Snyder, 2015). In this paper, we primarily focus 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 negative 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—such a benchmark is seldom observable and may even be impossible to define—and instead 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 media slant.

Baseline Measure. In the GKG 2.0 database, GDELT's baseline article tone comprises three metrics: Positive Score, Negative Score, and Tone. Specifically, for a given article q, Positive Score$_q$ is the percentage of words that carry a positive emotional connotation:

$$\text{Positive Score}_q = \left(\frac{\text{Number of positive words}_q}{\text{Total number of words}_q} \right) \times 100 \tag{1}$$

where q indexes the news article. In contrast, the $Negative\ Score_q$ is the percentage of all words in the article that carry a negative emotional connotation:

$$\text{Negative Score}_q = \left(\frac{\text{Number of negative words}_q}{\text{Total number of words}_q} \right) \times 100 \tag{2}$$

The $Tone_q$ is an aggregate measure representing the article's overall sentiment, derived from the difference between the Positive Score and the Negative Score:

$$\text{Tone}_q = \text{Positive Score}_q - \text{Negative Score}_q \tag{3}$$

Accordingly, the tone of a news article in GKG 2.0 database is quantified as a score ranging from -100 (extremely negative) to +100 (extremely positive), with common values falling between -10 and +10, and 0 indicating a neutral tone.

Validation. To validate the GKG 2.0 tone measures, we randomly sample 50,000 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 English-language 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 A1 of the Online Appendix, we compare the tone measures we derived with those provided in the GKG 2.0 database and find significant Pearson correlations ranging from 0.66 to 0.81. These results suggest that the tone measures in the GKG 2.0 database are both informative and valid.[9] The tone measures in the GKG 2.0 database capture sentiment information comparable to widely used algorithms in the literature. Additionally, GKG 2.0 offers alternative tone measures based on the Lexicoder Sentiment Dictionary (LSD) and SentiWordNet 3.0 (SWN). Later in this study, we 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.

$$\text{Tone}_{i,j,s,m} = \frac{\sum_{q \in \Omega_{i,j,s,m}} \text{Tone}_q}{\text{N}_{i,j,s,m}} \tag{4}$$

where $\text{Tone}_{i,j,s,m}$ denotes the average tone for news articles published by media outlet i, covering theme s, and referencing events in country j during month m.

Summary Statistics and Descriptive Patterns. The summary statistics of the baseline sample are presented in Panel A of Table 1. The mean and the median of the tone measure are -1.925 and -1.673, respectively. The sentiment variable demonstrates a large variation with a standard deviation of 3.754. On average, the tone is negative, reflecting the higher proportion of negative words relative to positive words.

Figure 4 presents distributional properties of the tone variable. Panel (a) plots the densities of positive and negative tone values, while Panel (b) shows the density of the tone measure, suggesting that the tone measures exhibit approximately normal statistical properties. Panel (c) plots the mean and corresponding 95% confidence intervals over time. Given the high granularity of the baseline sample, the tone measure shows substantial cross-sectional variation and does not exhibit a clear intertemporal pattern.

Figure 5 illustrates how average tone varies and evolves across prominent source–target

[9] 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 the website for variable construction. This practice inevitably introduces some degree of noise.

country pairs. Panel (a) shows that U.S. outlets portray Russia far more negatively than either the United Kingdom or China. Conversely, Panel (c) reveals that Russian media coverage of the United States is also consistently negative. For both U.S. and Russian coverage, sentiment toward China drops sharply with the onset of the COVID-19 pandemic in 2020. 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.[10]

Figure 6 focuses on the Russia–Ukraine dyad. Media tone in Ukrainian outlets toward Russia dropped precipitously following the onset of the Russia-Ukraine war in early 2022. In contrast, Russian media sentiment toward Ukraine declined more modestly over the same period.[11] 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 capture the institutional framework of a country, we employ the liberal democracy index from the Varieties of Democracy (V-Dem) project (Coppedge et al., 2023). Each indicator within the V-Dem dataset is assessed by at least five country experts who independently rate the indicators (Pemstein et al., 2022). The liberal democracy index provides an assessment of the extent to which a country adheres to the principles of liberal democracy, including the rule of law, checks and balances, and civil liberties. This index operates on a scale from 0 to 1, with higher values indicating a high degree of liberal democracy.

Let *Liberal Democracy Index*$_{A,t}$ and *Liberal Democracy Index*$_{B,t}$ represent the liberal democracy indices of country A and country B in year t, respectively. We define the institutional distance between country A and country B in year t, *Institutional Distance*$_{A,B,t}$ as:

$$\text{Institutional Distance}_{A,B,t} = \left|\text{Liberal Democracy Index}_{A,t} - \text{Liberal Democracy Index}_{B,t}\right| \tag{5}$$

The variable *Institutional Distance*$_{A,B,t}$ is our main independent variable. Panel B of Table 1 reports its summary statistics, and Figure 7 explores its distribution. As shown in Panel (a) of Figure 7, institutional distance varies widely across country pairs. Most country pairs have relatively small institutional distances, while their frequency declines as distance

[10] The tone measures in Figure 5 are residualized with respect to theme-by-year-month fixed effects to account for variation in topic composition over time. Patterns are similar without such residualization.

[11] These dynamics echo asymmetric public opinion reactions observed during earlier stages of the conflict (Korovkin and Makarin, 2023). See also Abramenko, Korovkin, and Makarin (2024) on the sharp shifts in the language of Ukraine's social media content before and after the full-scale invasion.

increases. Panel (b) indicates that, although institutional distance is not highly volatile, there is still meaningful time-specific variation once we residualize 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.

Additionally, we construct alternative measures of institutional distance by using the electoral democracy index from the V-Dem project and the Polity2 score 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 account broader dimensions of democracy, such as civil liberties and checks and balances, providing a more comprehensive assessment of political institutions. It is also important to note, however, that the Polity2 score, while widely used, is only updated through 2018. As part of our robustness checks, we confirm that our results remain consistent when using these two alternative democracy measures.

2.4 Other Data

We retrieve datasets from the following 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 (Bailey, Strezhnev, and Voeten, 2017; Voeten, 2000, 2013). The UN voting data is 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). Additionally, we 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. Moreover, 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 1.

Country-year level variables. To investigate the effect heterogeneity, 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." It is important to note that the first two indicators, "government censorship effort" and "internet censorship effort," assign higher values to indicate lower levels of censorship in the V-Dem database. To avoid potential confusion, we choose to rename these two indicators as "traditional media freedom" and "internet freedom," respectively. The summary statistics of these variables are presented in Panel C of Table 1.

Firm-specific variables. To examine the economic relevance of global media coverage, we analyze how financial markets respond to intensified media biases across countries. Financial data provide two key advantages for this purpose: first, they allow for the construction of granular firm-level exposure measures; second, market prices react quickly to relevant information, enabling clean identification of short-term responses to shifts in media tone. For the equity market analysis, we use data from the Center for Research in Security Prices (CRSP), Compustat, and the FactSet Revere dataset. We describe the construction of these variables and the corresponding empirical designs in Section 6.

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, we assume that their answers as a reflection of overall favorability or the broader "image" of the country.

3 Institutional Distance and News Coverage Tone

3.1 Baseline Specification

We begin by examining whether the tone of a country's media toward foreign nations varies systematically with institutional distance—an indicator of potential ideological or geopolitical bias. Following the logic of gravity models in international trade (Anderson and Van Wincoop, 2003), we estimate the following baseline specification:

$$\text{Tone}_{i(c),j,s,m(t)} = \beta \, \text{Institutional Distance}_{c,j,t-1} + \theta_{j,s,t} + \delta_{i(c),s,t} + \epsilon_{i(c),j,s,m(t)} \tag{6}$$

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 $\text{Institutional Distance}_{c,j,t-1}$, defined as the absolute difference in liberal democracy scores between countries c and j in year $t-1$, drawn from the Varieties of Democracy (V-Dem) dataset (for details, see Section 2.3).[12]

The richness of the dataset allows us to include two high-dimensional sets of fixed effects, mirroring gravity models in the trade literature (Larch and Yotov, 2024). $\theta_{j,s,t}$ absorbs 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

[12] Using a lagged measure helps ensure that our estimates are not mechanically affected by contemporaneous media coverage, though results are similar when using contemporaneous institutional distance.

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 a proxy for geopolitical and ideological misalignment. We further demonstrate 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. If such a robust association emerges, it would already provide strong evidence of global media bias along ideological and geopolitical lines. However, to further bolster a causal interpretation, we complement the baseline specification with a difference-in-differences design that exploits episodes of democratic erosion in the source country (see Section 3.4).

3.2 Baseline Results

Table 2 presents the baseline results. Across columns (1)–(3), we successively 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 (6), 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 6. 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. Negative sentiments often play a crucial role in propaganda, particularly in the portrayal of hostile countries. Specifically, negative sentiments against 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, making internal conflicts appear less significant in comparison. As supporting evidence, we disaggregate the tone variable into positive and negative components. The results in Online Appendix Table A2 indicate that negative sentiment is more strongly associated with institutional distance than positive sentiment, even after accounting for differences in their respective distributions.

We also draw on fine-grained sentiment measures from GDELT, which classifies emotion using the WordNet-Affect (WNA) lexicon (Strapparava and Valitutti, 2004). This resource supplies detailed affective annotations, allowing us to capture nuanced emotional content in news text. For each article, we compute WNA scores as the share of words linked to each emotional category, then aggregate these measures at the outlet–target-country–theme–month level. Figure 8 reports estimates from our baseline specification (6), now linking institutional distance to a range of WNA sentiment categories (full results in Online Appendix Table A3). 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.[13]

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. To address this, Panel A of Table 3 examines media bias using an alternative measure of geopolitical alignment: ideal point distance in UNGA voting decisions (Bailey, Strezhnev, and Voeten, 2017; Voeten, 2000, 2013). 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 other time-invariant indicators of bilateral similarity. Specifically, we add controls for shared official language (Melitz and Toubal, 2014), religious distance (Dow and Karunaratna, 2006), genetic distance (Spolaore and Wacziarg, 2009), and geographic distance. 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.[14] These findings suggest that insti-

[13] We validate the WNA sentiment measures through Pearson correlations with our baseline tone measures; see Online Appendix Figure A2.

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

tutional distance captures distinct and substantively meaningful aspects of ideological and political divergence, which are particularly salient in shaping international media tone.

Robustness. We conduct an extensive battery of robustness checks.

First, we examine whether our results are sensitive to alternative model specifications. In columns (1) and (2) of Online Appendix Table A5, we introduce country-pair fixed effects. The inclusion of country-pair fixed effects absorbs all time-invariant confounders at the country-pair level, including the time-invariant distance measures discussed above. 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, the findings are robust to alternative standard error clustering methods, as shown in Online Appendix Table A6. In the baseline estimation, we cluster the standard errors at the country-pair level. We additionally show that coefficients are consistently significant when clustering at media outlet, or source country level. Following Petersen (2008), we also employ 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.

Third, we consider alternative measures of both the dependent and independent variables used in our baseline analysis. Online Appendix Table A7 presents results using two alternative tone measures based on Lexicoder Sentiment Dictionary (LSD; Young and Soroka, 2012) and SentiWordNet 3.0 (SWN; Baccianella, Esuli, and Sebastiani, 2010) provided in the GKG 2.0 database. The results are consistent with our baseline findings. Online Appendix 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 Polity2 score. Importantly, the effect sizes, relative to the standard deviation of the outcome variable, are quite similar across all measures. Another measurement-related concern is that sentiment measures derived from GDELT's machine-translated text may be noisy, as bias could be associated with the quality of the translation. We randomly sample 50,000 articles per month from the GDELT GKG 2.0 database to construct an article-level dataset, and 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 construct a variable capturing the proportion of English-language articles and find that the estimated coefficient on institutional distance remains stable when controlling for the proportion of

English articles. Furthermore, the interaction between the share of English articles and institutional distance is not statistically significant. We further restrict 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 Online Appendix Table A9. We also iteratively exclude each major language from the sample and find that the results remain largely unaffected (Online Appendix Figure A3).

Fourth, we employ a range of alternative rules to identify each outlet's source country. Online Appendix Figure A4 (detailed in 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 look-ups, ChatGPT classifications, or based on top-level domain suffixes (e.g., ".uk" for the United Kingdom). Limiting the sample to roughly 180 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 towards institutionally distant countries, further reinforcing our baseline findings.

Fifth, we conduct additional tests using alternative samples. We report the results based on the tone variable winsorized at 0.5%, 1%, and 5%, respectively, in Online Appendix 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 perform a robustness check by excluding all observations related to the United States. The results, reported in Online Appendix Table A12, are quantitatively similar to our baseline estimates. Additionally, Online Appendix Figure A5 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. In addition, we aggregate the baseline sample to the country-pair–by–theme–by–year level, and the results remain consistent with our main findings (Online Appendix Table A14).

We also replicate the baseline result presented in Table 2 by constructing an alternative sample using the GDELT Event 2.0 database. In Online Appendix B, we provide a detailed discussion on 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 A13. Unlike the baseline specifications, this analysis incorporates Event Type fixed effects instead of Theme fixed effects, as the GDELT Event 2.0 database 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.

Finally, we further control for economic interdependence and conflicts between the source and target countries. Specifically, in Online Appendix 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.

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 A4, the degree of media bias increases over time. The coefficient grows annually by 12% (=0.023/0.191 as in column (2)) relatively to the baseline level.

Next, we estimate a model interacting institutional distance with yearly indicators:

$$\begin{aligned}\text{Tone}_{i(c),j,s,m(t)} = &\sum_{t=2015}^{2022} \beta_t(\text{D}_t \times \text{Institutional Distance}_{c,j,t-1}) \\ &+ \delta_{i(c),s,t} + \theta_{j,s,t} + \epsilon_{i(c),j,s,m(t)}\end{aligned} \tag{7}$$

where D_t is an indicator for year t and other terms follow the definitions in Equation (6). Figure 9 plots the year-specific coefficients β_t. 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.

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 estimate the following model:

$$\begin{aligned}\text{Tone}_{i(c),j,s,m(t)} = &\sum_{s\in\Omega^s} \beta_s(\text{W}_s \times \text{Institutional Distance}_{c,j,t-1}) \\ &+ \delta_{i(c),s,t} + \theta_{j,s,t} + \epsilon_{i(c),j,s,m(t)}\end{aligned} \tag{8}$$

where W_s is an indicator variable set to one for theme s and zero otherwise. Other model specifications follow the definitions in Equation (6). We plot the estimated the coefficients of

β_s in Figure 10 and report the detailed results in Online Appendix Table A16. 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."

3.3 Biased News Coverage on National Leaders

Although the breadth of GKG 2.0 article themes is a strength—allowing us to demonstrate that ideologically driven media bias extends beyond overtly political topics—one might expect stronger evidence of bias for explicitly political coverage. To test this, we focus on news articles about national leaders, who personify their countries on the global stage and are closely associated by the public with their regimes' political actions (e.g., Saich, 2014).

Using a curated sample of leader-related news articles (see Online Appendix B.2 for the details), we report results in Table 4, estimating a series of increasingly granular fixed effects specifications. In column (1), we include outlet, leader, and year fixed effects, isolating variation in how different outlets cover the same leaders while accounting for time-invariant outlet traits and common yearly shocks. Column (2) adds outlet $\times$ year and leader $\times$ year fixed effects, controlling for annual shifts in both outlet behavior and leader visibility. In column (3), we employ our most demanding specification with outlet $\times$ year-month and leader $\times$ year-month fixed effects, allowing us to identify media bias by comparing how different outlets cover identical leaders within the same month, while controlling for general outlet-specific monthly changes in coverage tone.

Across all specifications, the coefficient on institutional distance is both statistically significant and large—more than three times the baseline effect in Table 2. In column (2) of Table 4, increasing institutional distance from zero to one is associated with a 0.916-point drop in monthly average tone, equal to 47.4% of the full-sample mean and roughly two-thirds of the decline seen in Ukraine-based outlets covering Russia after the 2022 invasion (Figure 6). These findings underscore that media bias is especially pronounced in highly salient political contexts, such as coverage of national leaders.

3.4 Democratic Erosion as Natural Experiments

We exploit episodes of democratic erosion (or autocratization) in the source country to provide further evidence that the institutional distance between the source country and the target country plays a critical role in shaping global media bias.

Drawing on the Regimes of the World (RoW) classification from the V-Dem dataset (Lührmann, Tannenberg, and Lindberg, 2018), we define episodes of democratic erosion (or autocratization) as instances where a country experiences a downward shift in the degree of democracy: transitioning from a liberal democracy to an electoral democracy, from an electoral democracy to an electoral autocracy, or from an electoral autocracy to a closed autocracy. To account for early signs of democratic backsliding prior to formal regime type reclassification, we leverage the Episodes of Regime Transformation (ERT) dataset (Maerz et al., 2023) to identify the source countries where democratic erosion was initiated before 2015, the starting year of our sample. We exclude these cases, yielding 17 democratic erosion episodes in the sample. In addition, to focus on democratic erosion, we exclude source countries that underwent democratic transitions during the sample period. A list of treated countries is presented in Online Appendix Table A17.

If institutional distance—and the ideological and political divergence it captures—truly drives media bias, then a country's democratic backsliding should widen its distance from democracies and lead its media outlets to portray democratic countries more negatively. We test this implication with a continuous difference-in-differences (DID) design that estimates how democratic erosion in the source country changes the tone of its coverage of more- versus less-democratic target countries. The model is specified as follows:

$$\begin{aligned}\text{Tone}_{i,j,s,m} =& \beta\, \text{Erosion}_{c,t} \times \text{Liberal Democracy Index}_{j,t}^{\text{Target Country}} \\ &+ \delta_{i,s,t} + \theta_{j,s,t} + \alpha_{i,j} + \epsilon_{i,j,s,m}\end{aligned} \tag{9}$$

where $\text{Erosion}_{c,t}$ is an indicator variable that equals one for the years following democratic erosion in country c; $\text{Liberal Democracy Index}_{j,t}^{\text{Target Country}}$ represents the degree of democracy in the target country j in year t; and β captures the change in media tone toward countries with higher levels of democracy (relative to those with lower levels of democracy) after the source country experiences democratic erosion.

We further explore the dynamic effects using the following model specification:

$$\begin{aligned}\text{Tone}_{i,j,s,m} =& \sum_{\substack{\tau=-4 \\ \tau\neq -1}}^{4} \beta_\tau \left(\text{Erosion}_{c,\tau} \times \text{Liberal Democracy Index}_{j,\tau}^{\text{Target Country}}\right) \\ &+ \delta_{i,s,t} + \theta_{j,s,t} + \alpha_{i,j} + \epsilon_{i,j,s,m}\end{aligned} \tag{10}$$

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.

To probe robustness, we also replace the continuous measure of democracy in the target

country from Equations (9) and (10) with a binary indicator, which equals one if the target country is classified as a liberal democracy. To avoid the negative weights issues associated with staggered DID designs, we estimate the dynamic effects using the Sun and Abraham (2021) estimator and also report OLS estimation results for comparison.

Online Appendix Table A18 presents estimates for Equation (9). As suggested by column (2), democratic erosion results in a 0.089 decrease in media tone toward democratic countries compared to non-democratic countries, which corresponds to 2.38% of the standard deviation of the outcome variable (3.747 as reported in Table 1). The dynamic estimates are visualized in Figure 12, where Panel (a) displays the results using the continuous measure of democracy in the target country, and Panel (b) reports the results using the binary democracy indicator. The complete set of results is provided in Online Appendix Table A19. To assess the robustness of our results to individual episodes of democratic erosion, we re-estimate the model specification from column (2) of Table A18, excluding one episode at a time. The results, presented in Online Appendix A6, remain consistent and robust. Overall, the estimated coefficients for the interaction between the target country's democratic status and event-year dummies exhibit an increasingly downward trend, increase in magnitude, and attain statistical significance following democratic erosion. Importantly, we find no evidence of pre-existing trends, reinforcing the validity of our identification strategy. These findings further support the role of institutional distance in shaping global media bias.

3.5 Heterogeneity and Mechanisms

Thus far, our results point to a consistent pattern of global media bias: even after conditioning on an extensive set of fixed effects and controls, countries systematically cover their more institutionally distant counterparts in a more negative light. This relationship proves highly robust and is further corroborated by staggered DID estimates around episodes of democratic erosion. In this section, we explore several dimensions of heterogeneity that may shed light on the mechanisms behind this relationship. Specifically, we examine: (i) whether the pattern is driven mainly by liberal democracies or by less-democratic regimes; (ii) whether domestic controls on information flows intensify the bias; and (iii) whether economic or political stress at home is linked to more negative coverage of foreign countries.

3.5.1 Democracy in Source and Target Countries

We begin by examining whether the extent of media bias differs according to the source country's regime. To do so, we augment our baseline specification (Equation (6)) with an interaction between institutional distance and an indicator for whether the source country

is classified as a liberal democracy:

$$\begin{aligned}\text{Tone}_{i(c),j,s,m(t)} = \beta\, &\text{Institutional Distance}_{c,j,t-1} \\ &+ \lambda\, \text{Institutional Distance}_{c,j,t-1} \times \text{Liberal Democracy Indicator}_{c,t-1}^{\text{Source Country}} \\ &+ \delta_{i(c),s,t} + \theta_{j,s,t} + \epsilon_{i(c),j,s,m(t)}\end{aligned} \tag{11}$$

where $\text{Liberal Democracy Indicator}_{c,t-1}^{\text{Source Country}}$ is an indicator variable that equals one when the source country is a liberal democracy in year $t-1$ according to the Regimes of the World (RoW) measure from the V-Dem dataset (Lührmann, Tannenberg, and Lindberg, 2018).

Likewise, we use the following model specification to study the heterogeneity of our baseline estimates by whether the target country is the liberal democracy:

$$\begin{aligned}\text{Tone}_{i(c),j,s,m(t)} = \beta\, &\text{Institutional Distance}_{c,j,t-1} \\ &+ \lambda\, \text{Institutional Distance}_{c,j,t-1} \times \text{Liberal Democracy Indicator}_{j,t-1}^{\text{Target Country}} \\ &+ \delta_{i(c),s,t} + \theta_{j,s,t} + \epsilon_{i(c),j,s,m(t)}\end{aligned} \tag{12}$$

where $\text{Liberal Democracy Indicator}_{j,t-1}^{\text{Target Country}}$ is an indicator variable that equals one when the target country is a liberal democracy in year $t-1$.

In columns (1)–(2) of Table 5, we examine how the degree of political-regime-driven media bias varies depending on whether the source country is a liberal democracy. Column (1) includes outlet × year fixed effects and target country × theme × year fixed effects; column (2) adopts the most detailed baseline specification with outlet × theme × year fixed effects and target × 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 2).

Figure 11 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 essentially flat for liberal democracies.

Columns (3) and (4) of Table 5 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 (–0.466 in column (4) versus –0.260 in Table 2); for less democratic targets it is smaller and only marginally significant (–0.106). Panels (c) and (d) of Figure 11 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 depend not only on the source country's regime but also, and even more strongly, on whether the country being covered is democratic.

3.5.2 Information Control

To sustain regime stability, governments—particularly authoritarian ones—frequently 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).

In this section, we examine how the extent of global media bias varies across countries with different levels of information control. We estimate the following heterogeneity specification:

$$\begin{aligned}\text{Tone}_{i(c),j,s,m(t)} = \alpha\, &\text{Institutional Distance}_{c,j,t-1} + \beta\, \text{Institutional Distance}_{c,j,t-1} \times \text{Z}_{c,t-1} \\ &+ \delta_{i(c),s,t} + \theta_{j,s,t} + \epsilon_{i(c),j,s,m(t)}\end{aligned} \tag{13}$$

where $Z_{c,t-1}$ indicates the levels of information control in source country c in year $t-1$. We use different indicators to capture the levels of information control as detailed below.

Media Freedom. Panel A of Table 6 explores whether the association between institutional distance and tone differs with the level of media freedom in the source country. Columns (1)–(2) interact institutional distance with a measure of traditional-media freedom (TV, radio, print), while columns (3)–(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 (i.e., near-total censorship)—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 B turns to media ownership. Columns (1)–(2) interact institutional distance with a measure of state ownership in print media; columns (3)–(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.

Online Appendix Table A20 offers a complementary test using a binary indicator for whether an outlet itself is state-owned (classified via GPT-4o API). The baseline coefficient for non-state outlets is –0.23, while the interaction is about –0.29, suggesting a noticeably larger association for state outlets. Taken together, these patterns suggest that state media ownership in the source country is associated with a relatively more negative coverage toward institutionally distant countries, consistent with stronger ideological and political media bias.

3.5.3 Economic Distress

We next examine whether the association between institutional distance and media tone is 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 7.

Columns (1)–(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 towards more institutionally distant countries. Columns (3)–(4) replace growth with an indicator for high inflation (>5 percent). 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

In addition to economic challenges, we further investigate whether domestic instability prompts regimes to strategically shape media narratives by portraying institutionally distant countries unfavorably. 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 Database 2.0. This index measures people's resentment towards their government by calculating the proportion of events with negative sentiment scores among all events targeted at the government.

Column (1) of Table 8 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 sources 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.

In sum, the heterogeneity patterns point to a consistent logic: media bias is most pronounced when the *source* country is an autocracy and the *target* is a liberal democracy, and it intensifies under conditions that tighten the government's grip on information or heighten domestic economic and political pressure at home. These patterns are consistent with the idea that authoritarian governments strategically lean on outlets controlled or influenced by the state to cast democracies in a negative light, both to showcase their own model as superior and to deflect blame for economic or political troubles. During the COVID-19 pandemic, for example, Chinese state media criticized the United States for its pandemic response, highlighting the US government's mismanagement and lack of accountability, while praising China's capacity for rapid social mobilization (Fu, 2023; Xia, Huang, and Zhang, 2022).[15] Moreover, authoritarian regimes frequently deflect criticism by attributing internal

[15]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); Global Times. April 9, 2020. "The US Has Dragged Down the Fight Against the Epidemic, and It Has Also

issues to external factors, as exemplified by Syrian state media framing the Arab Spring uprisings as foreign plots (Alrababa'h and Blaydes, 2021). Such anti-foreign messaging stirs nationalistic feelings, fostering a unifying "rally-round-the-flag" effect (Weiss, 2014). Therefore, biased reporting targeting democracies may serve as a strategic tool to divert attention from domestic issues, shore up popular support, and preserve authoritarian rule.

4 Institutional Distance and News Theme Selection

Media bias can arise not only from systematic differences in tone but also from selective topic choice, where outlets highlight themes that typically carry negative connotations. For example, Chinese media might emphasize incidents of gun violence when reporting on the United States. To test this empirically, we estimate the following model:

$$\begin{aligned} \text{Theme Share}_{c,j,s,t} = \beta \left(\text{Institutional Distance}_{c,j,t-1} \times \overline{\text{Theme Tone}}_s\right) \\ + \gamma \, \text{Institutional Distance}_{c,j,t-1} \\ + \delta_{j,s} + \theta_c + \alpha_t + \epsilon_{c,j,s,t} \end{aligned} \tag{14}$$

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 some specifications, we use a binary variable as the dependent variable to indicate whether news on theme s concerning target country j was reported by source country c in year t. Given the computational constraints, we aggregate the data at the source-country–target-country–theme–year level.

$\overline{\text{Theme Tone}}_s$ is the average tone of theme s and is measured in the following two ways. The first approach calculates the average theme tone with the average tone of all news articles covering theme s. This is expressed by the following equation:

$$\overline{\text{Theme Tone}}_s = \frac{1}{Q} \sum_{q=1}^{Q} \text{Tone}_{q(s)} \tag{15}$$

The other approach calculates the average tone for a specific target country j, theme s, and year t by averaging across tones of all related news articles while excluding articles published by media outlets from source country c:

$$\overline{\text{Theme Tone}}_{\substack{j,s,t \\ i \notin c}} = \frac{1}{Q} \sum_{q=1}^{Q} \text{Tone}_{\substack{q(j,s,t) \\ i \notin c}} \tag{16}$$

Dimmed the Prospects of the World," https://opinion.huanqiu.com/article/3xlaqBQqzYD (accessed May 21, 2025).

We expect the estimated coefficient β in Equation (14) to be negative, which would suggest that media outlets in a source country devote more space to inherently negative themes when covering countries that are institutionally distant. In other words, when a theme typically carries negative connotations (i.e., has a negative average tone), the stronger the positive relationship between institutional distance and that theme's share of total coverage.

Columns (1) and (2) of Table 9, which rely on two alternative measures of a theme's average tone, confirm the prediction: the interaction between institutional distance and theme tone is negative and statistically significant. Columns (3) and (4), which employ a binary dependent variable for whether a certain theme is at all covered, yield consistent results. Finally, column (5) zooms in on the "violence" theme, one of the more negative topics in our sample, and shows that greater institutional distance significantly raises the share of violence-related coverage. Taken together, these findings indicate that as institutional distance grows, outlets systematically emphasize negative themes when reporting on foreign countries. Hence, ideologically driven global media bias manifests not only in the tone of news coverage but also in topic selection.

5 Alliances in Ideological Warfare

Our findings so far show that media outlets portray institutionally similar countries more favorably. We now broaden our analysis and use media narratives to reveal *alliances* in a wider ideological contest among competing political-economic blocs. To do so, we devise a correlation-based alignment metric for each country pair, capturing how similarly their outlets report on *third* countries across many themes. The resulting measure is granular, high-frequency, and scalable, providing a novel estimate of ideological proximity in line with recent calls for finer "geoeconomic" indicators (Mohr and Trebesch, 2024).

To this end, we start with the tone variable used in our baseline estimation. We aggregate the tone measure at the source country c, target country j, theme s, and year t level:

$$\underset{j \neq c}{\text{Tone}_{c,j,s,t}} = \frac{\sum_{(i,j,s,m) \in \Omega_{c,j,s,t}} \text{Tone}_{i(c),j,s,m(t)}}{\text{N}_{c,j,s,t}} \tag{17}$$

where $Tone_{i(c),j,s,m(t)}$ is the average tone of media outlet i from country c when reporting country j for theme s during month m of year t. To ensure the stability of the Tone measure, we require that each combination of c, j, s, and t contains at least 12 observations, which means there should be reporting at least once per month.

Our goal is to gauge the extent to which media coverage co-moves between country A and country B, using their respective tone measures toward themes in countries other than

each other. Formally, for country A in year t, we have the tone measure $\text{Tone}_{\substack{A,j,s,t \\ j\neq A,B}}$ that capture its media sentiment toward theme s in country j, excluding any news about country A or B. Similarly, for country B, we define $\text{Tone}_{\substack{B,j,s,t \\ j\neq A,B}}$. We then use these two constructs to compare how countries A and B report on the same theme s in country j. That is, for each pair of countries (A, B) in year t, we construct the following Pearson correlation coefficient:

$$\rho_{A,B,t} = \frac{\text{Cov}(\text{Tone}_{A,j,s,t}, \text{Tone}_{B,j,s,t})_{j\neq A,B}}{\sigma_{A,t}\sigma_{B,t}} \tag{18}$$

where the numerator is the covariance between $\text{Tone}_{A,j,s,t}$ and $\text{Tone}_{B,j,s,t}$, and $\sigma_{A,t}$ and $\sigma_{B,t}$ are the standard deviations of $\text{Tone}_{A,j,s,t}$ and $\text{Tone}_{B,j,s,t}$, respectively, for year t. To ensure the reliability of the correlation coefficients, the sample used for their calculation must consist of at least 100 observations.

Panel (a) of Figure 13 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 "alliances."

The country-pair structure of $\rho_{A,B,t}$ allows us to conduct a network analysis. Panel (b) 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 correlation value between 2015 and 2022 falls in the top quintile. With information of regime types in 2015, Blue nodes denote liberal democracies; red nodes, less democratic countries. Liberal democracies cluster tightly around the United States at the center, whereas non-democracies form separate hubs, most prominently around Russia and, to a smaller extent, China. This structure mirrors the findings in Panel (a): in ideological contests, countries gravitate toward others with similar institutional characteristics and coordinate their media narratives accordingly.

In Figure 14, we explore the changing structure of alliance networks over time. Panel (a) displays changes in the average edge weight, a measure that quantifies the strength of interaction or connectivity between subnetworks of liberal and non-liberal democracies. Beginning in 2019, we observe a marked decline in connectivity between these groups, which may be plausibly linked to increasing geopolitical fragmentation or "decoupling" following the onset of the COVID-19 pandemic and the Russia–Ukraine war. Panel (b) presents changes in the clustering coefficient for liberal and non-liberal democracies, a measure that captures the degree of localized connectivity or internal cohesion within each group. Among liberal democracies, the clustering coefficient fluctuates modestly but remains relatively stable, averaging around 0.7 between 2015 and 2022. In contrast, the clustering coefficient for non-liberal democracies reaches a peak in 2017, followed by a consistent year-over-year

decline. This downward trend may reflect growing fragmentation or weakening alliance cohesion among non-liberal democracies. Panel (c) displays the evolution of degree centrality for the United States and the United Kingdom, both of which maintain consistently high and stable centrality levels in the liberal democracy network. In contrast, Panel (d) reveals a downward shift in the degree centrality of China starting in 2019, and of Russia in 2020, in the non-liberal democracy network. These trends may indicate a declining central role for both countries in coordinating or sustaining alliances among non-liberal democracies.

We also examine whether our novel measure of geopolitical and ideological alignment based on the correlations in media narratives about third countries has predictive power for other forms of international alignment. Online Appendix Table A21 shows that the media alliance measure between country i and country j in year $t-1$ significantly predicts UN voting alignment (Bailey, Strezhnev, and Voeten, 2017; Voeten, 2000, 2013), the formation of bilateral investment treaties (Alschner, Elsig, and Polanco, 2021), and participation in regional trade agreements (Egger and Larch, 2008) between these countries in year t. These findings suggest that our media-based measure is not merely a descriptive indicator of shared narratives, but also a forward-looking proxy for deeper geopolitical alignment. Its ability to predict alignment across both diplomatic and economic domains indicates that media narrative similarities reflect latent strategic affinities that can precede—and potentially shape—the development of formal institutional linkages between countries.

6 Economic and Political Relevance of Media Bias

6.1 Financial Market Reactions

Media bias can have significant economic consequences. Calomiris and Mamaysky (2019) show that the context and content of news articles, including topic-specific sentiment, predicts risk and return in stock markets in 51 developed and emerging economies. Gurun and Butler (2012) document the association between local media slant and firm equity values in the United States. Baloria and Heese (2018) demonstrate that media slant has a profound impact on firm behaviors such as information disclosure. However, the existing literature largely focuses on the domestic effects of media biases. Evidence on the economic relevance of *global* media bias remains scarce. One exception is Kim (2018) who provides evidence that Chinese newspapers are biased against foreign automakers, leading to disproportionate public attention on foreign car recalls and a greater negative impact on their sales.

We examine the relationship between intensified media bias across countries and its economic relevance using the immediate responses from the financial markets. The efficient

market hypothesis suggests that the equity market would quickly incorporate any relevant information into prices (Fama, 1970). In contrast, other outcomes such as trade flows or direct investment, may adjust more slowly. The richness of firm-level data also allows us to measure each firm's exposure to global media bias with precision and to gauge the markets' immediate response.

Using information originally from corporate disclosure, we construct a firm-level measure of exposure to foreign media bias based on the revenue share from a given foreign country. We only focus on US firms as the data on foreign sales are limited for firms in other countries. Following the standard practices in the asset pricing literature, we first construct a monthly panel of individual stock returns using data from the Center for Research in Security Prices (CRSP). We restrict the sample to US-headquartered firms using information from Compustat. To measure firms' exposure to foreign media sentiment, we construct:

$$\text{Tone Exposure}_{i,t} = \sum_{c=1}^{C} Tone_{c,t} \times Rev_{i,c,t} \tag{19}$$

where $Tone_{c,t}$ is the average tone of media outlets from a given country c toward the US in year-month t. Following previous studies (e.g., Huang et al., 2023), we use data from the FactSet Revere database to measure a firm's revenue exposure to a country. $Rev_{i,c,t}$ is firm i's revenue share from country c as reported in the Factset Revere database. A decrease in *Tone Exposure* indicates an intensified media bias toward the firm's countries of interest.

Merging media exposure with the return panel yields a firm-month-level sample comprising 2,734 distinct firms from 2015 to 2022. Observations with missing $Tone\ Exposure_{i,t}$, stemming from firms not recording any sales to any foreign country in a certain year, are dropped because such data cannot serve as a meaningful counterfactual for sentiment exposure. Therefore, we compare firms with sales from countries with relatively positive sentiment toward the US with ones with sales from countries with relatively negative sentiment. We estimate the following panel regression model using OLS:

$$\text{Returns}_{i,t+1} = \beta\, \text{Tone Exposure}_{i,t} + \delta_i + \theta_t + \epsilon_{i,t} \tag{20}$$

where $Tone\ Exposure_{i,t}$ is the previously defined exposure for for firm i in month t; δ_i denotes firm fixed effects; θ_m denotes year-month fixed effects. The standard errors are clustered at the firm level. Consistent with our baseline regression, we consider a lead-lag model specification.

As shown in Panel A of Table 10, we present the findings for the monthly raw returns (*Return*) and standard asset pricing models (Fama and French, 1992, 1993) including market

model adjusted returns (*Return (CAPM)*) and returns adjusted by the Fama-French three factor model (*Return (FF3)*). In this sample, *Tone Exposure* is on average negative with a standard deviation of 0.353. As shown in column (1) of Table 10, one standard deviation decrease in *Tone Exposure* is associated with a 0.275% drop in monthly returns, which translates to 3.30% drop in annualized returns. The results are robust to alternative asset pricing models.

We construct a measure at the foreign country level, where US firms generate sales to account for the general sentiment of the destination due to idiosyncratic reasons. One concern is that the captured media bias towards US reflects a common negative or positive sentiment driven by local trends. We define:

$$\text{Tone Source Country}_{i,t} = \sum_{c=1}^{C} Tone_{c,t}^{Source} \times Rev_{i,c,t} \tag{21}$$

where $Tone_{c,t}^{Source}$ is the average sentiment from media outlets outside country c toward country c in month t. We include this measure in Panel B of Table 10 and the results are unaffected. In addition, Panel C reports results using a lead specification, where we include $\textit{Tone Exposure}_{i,t+2}$, as a falsification test for anticipatory effects.

Taken together, the sizable impact on stock returns suggests that stock market participants anticipate firms with greater exposure to negative sentiment toward the United States to experience lower cash flows generated from those countries.

6.2 Public Opinion Reactions

In the final analysis, we explore the potential political consequences of global media bias, specifically its relevance for public opinion toward foreign countries. We leverage the Gallup World Poll Survey data to estimate whether average media tone of source county i toward target country j in year $t-1$ can predict the public opinion of country i toward country j in year t. We measure public opinion favorability using Gallup's reported approval rate of the target country's leadership, based on the survey 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, this measure serves as a meaningful proxy for overall sentiment in the source country, as leadership approval tends to reflect the broader image of the target country (Goldsmith, Horiuchi, and Matush, 2021).

Table 11 presents the results.[16] Column (1) controls for target-country and year fixed

[16]Specifically, we estimate the following equation:

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

effects; column (2) adds source-country fixed effects; column (3) includes the most stringent set of controls, adding target × year and source × year fixed effects. Across all specifications, average media tone is positively associated with future public approval of the target's leadership. The fully saturated model in column (3) implies that a one-standard-deviation increase in tone toward a foreign country is roughly associated with a percentage-point subsequent rise in the approval of that country's regime. These estimates suggest the possible influential role of domestic media 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 between 2015 and 2022, we document that, even after conditioning on a rich set of fixed effects and controls, we find 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.

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

We also document the economic and political relevance of cross-country news coverage. In the equity market, U.S.-listed firms with greater exposure to countries exhibiting negative media tone experience contemporaneous declines in stock prices. In the credit market, negative tone is associated with higher borrowing costs and lower deal volumes in syndicated lending. We also find that media tone is predictive of future international public opinion of foreign governments. These findings show that international media coverage contains information about international relations relevant for financial markets and highlight 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

where $Y_{i(c),j,t}$ is a binary indicator for whether respondent i (from country c) approves the leadership of the target country j in year t; $\overline{\text{Tone}}_{c,j,t-1}$ 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, such as gender, income, education level, and employment status. Standard errors are clustered at the country-pair level.

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 and high-frequency alternative that complements 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 for research at the intersection of media, international economics, and political economy.

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.

Alschner, Wolfgang, Manfred Elsig, and Rodrigo Polanco. 2021. "Introducing the Electronic Database of Investment Treaties (EDIT): The Genesis of a New Database and Its Use." *World Trade Review* 20 (1):73–94.

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.

Baloria, Vishal P. and Jonas Heese. 2018. "The Effects of Media Slant on Firm Behavior." *Journal of Financial Economics* 129 (1):184–202.

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.

Calomiris, Charles W and Harry Mamaysky. 2019. "How News and its Context Drive Risk and Returns around the World." *Journal of Financial Economics* 133 (2):299–336.

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, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Sandra Grahn, Allen Hicken, Katrin Kinzelbach, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Anja Neundorf, Pamela Paxton, Daniel Pemstein, Oskar Rydén, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundström, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 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.

Egger, Peter and Mario Larch. 2008. "Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis." *Journal of International Economics* 76 (2):384–399.

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.

Fama, Eugene F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance* 25 (2):383–417.

Fama, Eugene F. and Kenneth R. French. 1992. "The Cross-Section of Expected Stock Returns." *The Journal of Finance* 47 (2):427–465.

———. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33 (1):3–56.

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

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.

Gurun, Umit G. and Alexander W. Butler. 2012. "Don't Believe the Hype: Local Media Slant, Local Advertising, and Firm Value." *The Journal of Finance* 67 (2):561–598.

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.

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.

Huang, Yi, Chen Lin, Sibo Liu, and Heiwai Tang. 2023. "Trade Networks and Firm value: Evidence from the U.S.-China Trade War." *Journal of International Economics* 145:103811.

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

Kim, Sung Eun. 2018. "Media Bias Against Foreign Firms as a Veiled Trade Barrier: Evidence from Chinese Newspapers." *American Political Science Review* 112 (4):954–970.

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.

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, Kately Arriaga, Anne Blackburn, Daniel de Kadt, Bryn Farrington, Joe Felter, Caterina Hyneman, Alexandra Leal Silva, Talia Lorch, Jack McLean, Mateo Prada, Marie Tobia, and Michael Weintraub. 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.

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: Summary Statistics

Panel A: Variables at Domain - Target - 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	.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 (Polity2 Score)	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 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, and month levels. Panel B presents the independent variables at the country-pair level. Panel C reports a set of country-year level variables. The sample is constructed in Online Appendix B. The definitions of all variables are provided in Online Appendix E.

Table 2: 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
Adjusted R^2	0.263	0.278	0.291
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 presents the baseline estimation results documenting the presence of bias in the tone with which outlets cover institutional distant countries in the news. The dependent variable, *Monthly Average Tone*, measures cross-country media bias and is calculated as the monthly average tone using the data from GDELT GKG 2.0 database. The sample is at the media outlet, target country, theme, and month level. Standard errors in parentheses are clustered at the country pair level. Variables are defined in Online Appendix E. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 3: 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
Adjusted R-squared	0.264	0.279	0.292
Number of Country Pairs	13501	13501	13501

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
Adjusted R-squared	0.265	0.280	0.293
Number of Country Pairs	8944	8944	8944
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 presents the baseline estimation results documenting the presence of bias in the tone with which outlets cover geopolitically distant countries in the news. The dependent variable, *Monthly Average Tone*, measures cross-country media bias and is calculated as the monthly average tone using the data from GDELT GKG 2.0 database. The sample is at the media outlet, target country, theme, and month level. Standard errors in parentheses are clustered at the country pair level. Panel A shows the results using *UN Voting Distance* as the key independent variable. Panel B reports the results incorporating all distance measures. Variables are defined in Online Appendix E. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Bias News Coverage on 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
Adjusted R-squared	0.145	0.193	0.279
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 presents 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. Standard errors in parentheses are clustered at the country pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 5: 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 × Liberal Democracy Indicator (Source)	0.251***	0.246***		
	(0.080)	(0.078)		
Institutional Distance × Liberal Democracy Indicator (Target)			-0.298***	-0.318***
			(0.115)	(0.111)
Observations	81,288,434	81,288,434	81,288,434	81,288,434
Adjusted R-squared	0.278	0.291	0.278	0.291
Number of Country Pairs	14,379	14,379	14,379	14,379
Target Country × Theme × Year FEs	Yes	Yes	Yes	Yes
Outlet FEs	No	No	No	No
Outlet × Year FEs	Yes	No	Yes	No
Outlet × Theme × Year FEs	No	Yes	No	Yes

Notes: This table presents the heterogeneous association between institutional distance and media sentiment, conditional on the regime type of the source or target country. *Liberal Democracy Indicator* is set to one if the source country or target country is a liberal democratic country, and zero otherwise. Other table specifications follow the conventions in Table 2.

Table 6: Information Control

Panel A: 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
Adjusted R-squared	0.278	0.291	0.278	0.291
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

Panel B: 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
Adjusted R-squared	0.279	0.292	0.279	0.292
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 presents the heterogeneous association between institutional distance and media sentiment, conditional on the extent of information control at the country level. Panel A examines whether the magnitude of the institutional distance coefficient depends on the media freedom of the source country. Panel B investigates whether the coefficient magnitude is influenced by the media ownership structure of the source country. Other model specifications follow the conventions in Table 2.

Table 7: 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
Adjusted R-squared	0.289	0.302	0.279	0.292
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 presents the heterogeneous association between institutional distance and media sentiment, conditional on the source country's economic conditions. Other model specifications follow the conventions in Table 2.

Table 8: Domestic Turmoil

	All Sample		Non-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
Adjusted R-squared	0.291	0.291	0.239	0.333
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 presents the heterogeneous association between institutional distance and media sentiment, conditional on the domestic turmoil of the source country. Other model specifications follow the conventions in Table 2.

Table 9: 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
Adjusted R-squared	0.205	0.205	0.459	0.460	0.227
Target × Theme FE	Yes	Yes	Yes	Yes	Yes
Source Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Country Pair FE	No	No	No	No	No

Notes: This table illustrates the presence of media bias from the perspective of theme selection. We construct the sample at the source country, target country, theme, and year level with information 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 10: Equity Market Reactions

Panel A. Without Control	(1)	(2)	(3)
	Dependent Variable (t+1)		
	Return	Return (CAPM)	Return (FF3)
Tone Exposure (t)	0.0078***	0.0073***	0.0068***
	(0.003)	(0.003)	(0.003)
Observations	169,806	169,806	168,483
Adjusted R-squared	0.1992	0.0442	0.0177
Mean of Tone Exposure	-0.405	-0.405	-0.405
SD of Tone Exposure	0.353	0.353	0.353
Mean of Dependent Variable	0.005	-0.005	-0.004
SD of Dependent Variable	0.138	0.126	0.127
Firm FEs	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes

Panel B. Controlling for Source Country Sentiment	(1)	(2)	(3)
	Dependent Variable (t+1)		
	Return	Return (CAPM)	Return (FF3)
Tone Exposure (t)	0.0079***	0.0070***	0.0063**
	(0.003)	(0.003)	(0.003)
Tone Source Country (t)	-0.0003	0.0007	0.0013
	(0.002)	(0.002)	(0.002)
Observations	169,806	169,806	168,483
Adjusted R-squared	0.1992	0.0442	0.0177
Firm FEs	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes

Panel C. Falsification	(1)	(2)	(3)
	Dependent Variable (t+1)		
	Return	Return (CAPM)	Return (FF3)
Tone Exposure (t)	0.0073**	0.0073**	0.0078**
	(0.003)	(0.003)	(0.003)
Tone Exposure (t+2)	0.0017	0.0007	-0.0005
	(0.003)	(0.003)	(0.003)
Observations	167,906	167,906	166,599
Adjusted R-squared	0.1994	0.0442	0.0176
Firm FEs	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes

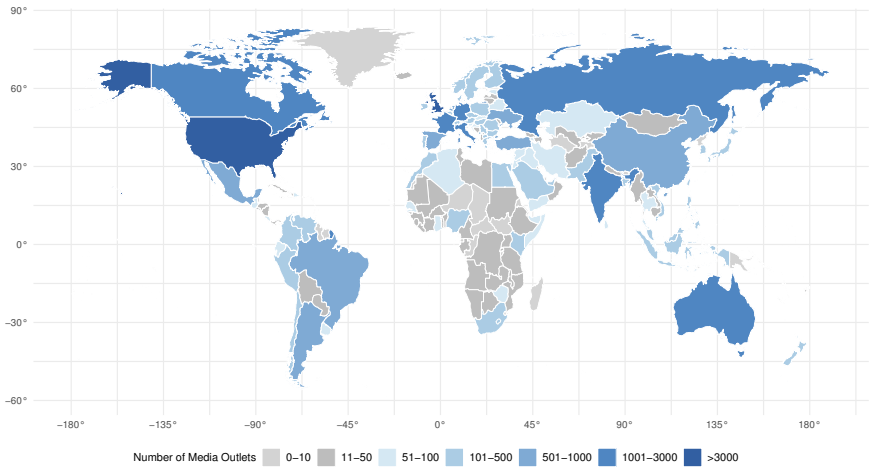
Notes: This table presents evidence on the equity market's response to global media bias. We use monthly stock return data for a sample of U.S. firms from the Center for Research in Security Prices (CRSP), merged with Compustat. In Panel A, we construct the variable: $Tone\ Exposure_{i,t} = \sum_{c=1}^{C} Tone_{c,t} \times Rev_{i,c,t}$, where $Rev_{i,c,t}$ denotes firm i's revenue share from country c in month t, sourced from the FactSet Revere Geographic Revenue database. $\text{Tone}_{c,t}$ represents the average sentiment (tone) of media outlets in country c toward the United States in month t. The final dataset is at the firm-month level, covering 2,734 distinct firms over the period 2015–2022. *Return* is the raw monthly return. *Return (CAPM)* is Returns adjusted using the Capital Asset Pricing Model. *Returns (FF3)* is the abnormal monthly return adjusted by the Fama-French 3 factor model. Dependent variables are measured in month t. In Panel B, $Tone\ Source\ Country_{i,t} = \sum_{c=1}^{C} Tone_{c,t}^{Source} \times Rev_{i,c,t}$, where $Tone_{c,t}^{Source}$ is the average sentiment from media outlets outside country c toward country c in month t. Panel C reports results using a lead specification, where we include $Tone\ Exposure_{i,t+2}$. All regressions include firm fixed effect and year-month fixed effects. Standard errors are clustered at the firm level and reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 11: Public Opinion Reactions

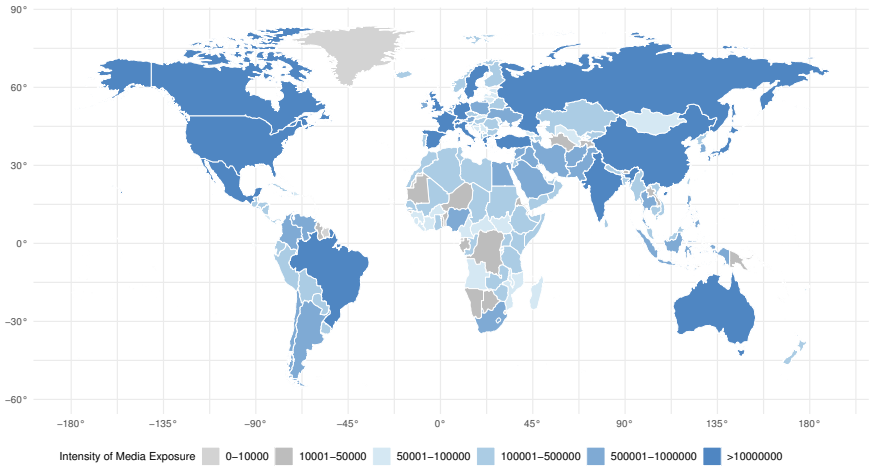
	Dependent Variable: Approval		
	(1)	(2)	(3)
Tone (Previous Year)	0.023***	0.008***	0.009**
	(0.005)	(0.003)	(0.004)
Observations	3,570,351	3,570,351	3,570,351
Adjusted R-squared	0.049	0.125	0.143
Mean of Approval	0.521	0.521	0.521
SD of Approval	0.500	0.500	0.500
Mean of Tone	-1.421	-1.421	-1.421
SD of Tone	1.263	1.263	1.263
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 is predictive of 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 to the target country in a given year. The dependent variable, *Approval*, equals 1 if the respondent approves the leadership performance of the target country and 0 otherwise. The independent variable is lagged by one year to mitigate simultaneity concerns. Standard errors are clustered at the source country level and reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Figures



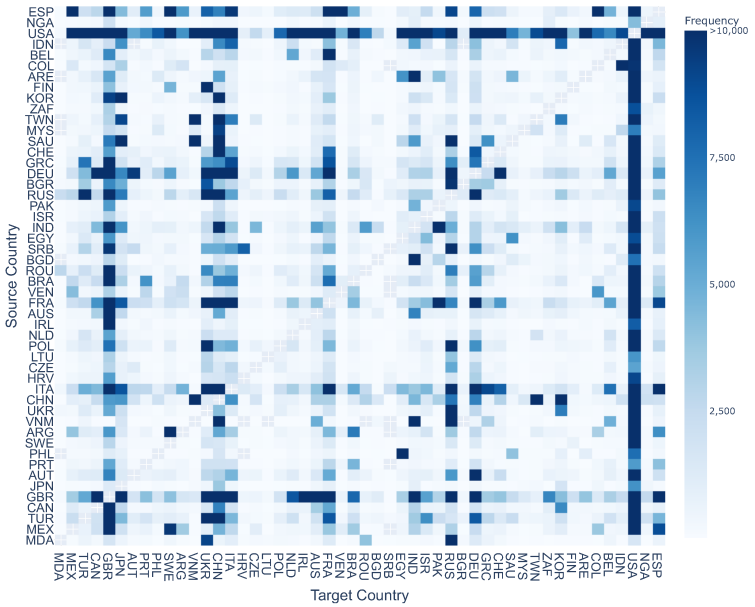
(a) Global Distribution of Media Outlets



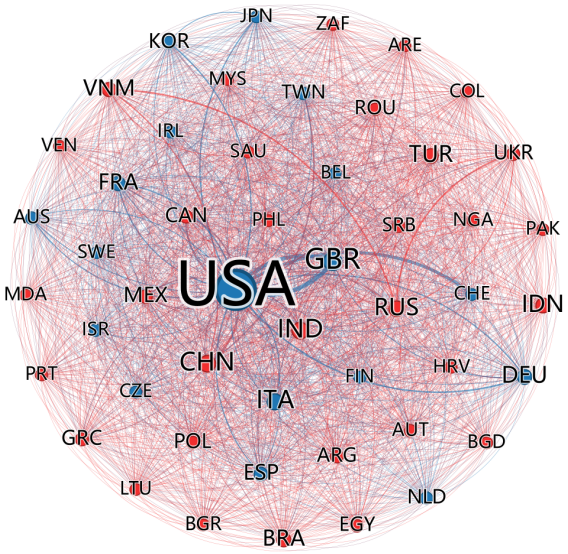
(b) Media Exposure

Figure 1: 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 coverage across countries, measured by the number of articles by theme covering each target country. If an article addresses multiple themes, it is counted once for each relevant theme, thereby contributing multiple times to the coverage intensity.



(a) Bilateral Frequency



(b) Reporting Network

Figure 2: Bilateral Reporting Relations

Notes: Panel (a) presents a heat map illustrating the intensity of news coverage from one country to another. For simplicity, we focus on the top 50 countries ranked by the total number of news articles issued by their media outlets in 2022. The Y-axis represents the source country; the X-axis represents the target country. Color intensity reflects the number of news articles by theme from media outlets in the source country that cover issues in the target country. Panel (b) displays a network graph where node size corresponds to the total number of news articles issued. Edge thickness represents the volume of coverage from one country to another. Green nodes denote liberal democracies; red nodes denote non-liberal-democratic countries.

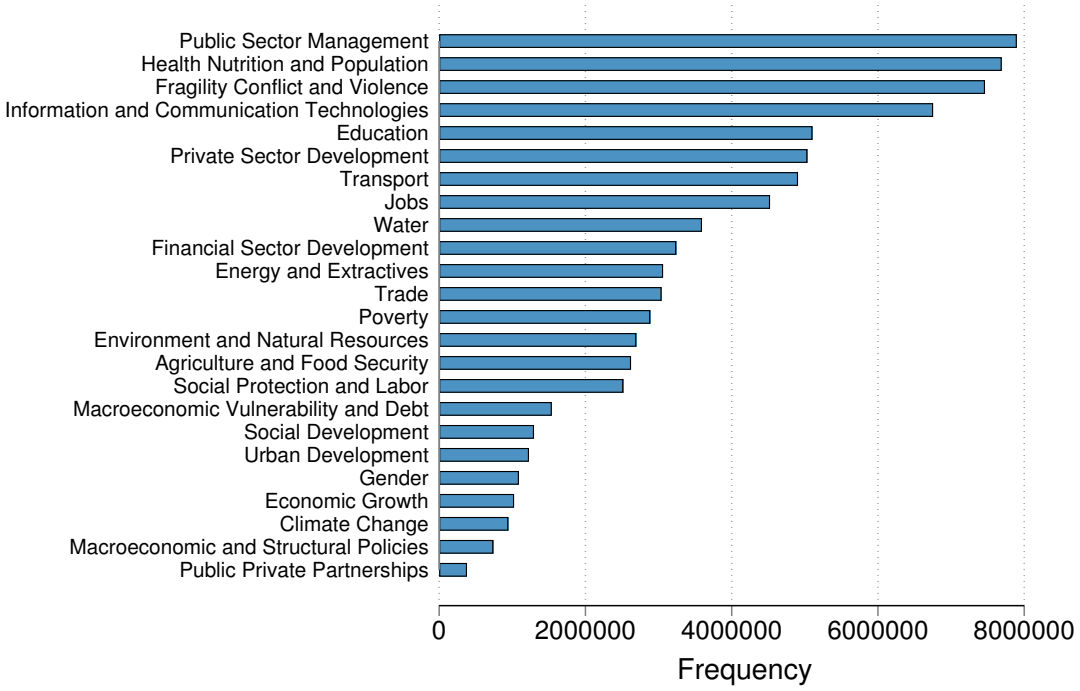
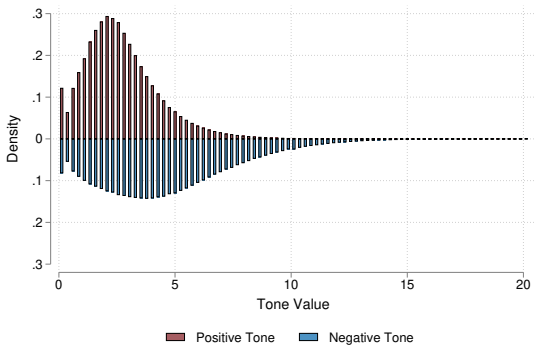
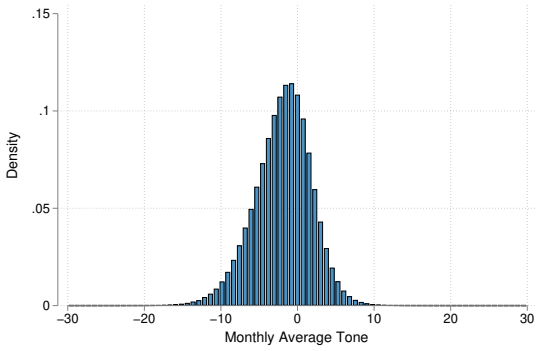


Figure 3: Themes Distribution

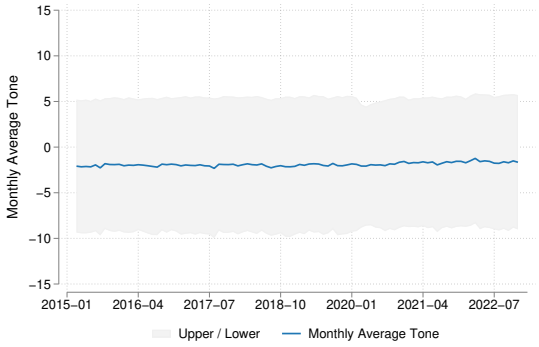
Notes: This figure displays the frequency of observations by theme in the baseline sample. Themes are classified according to Level 1 of the World Bank Group's Topical Taxonomy.



(a) Positive & Negative Tone Distribution



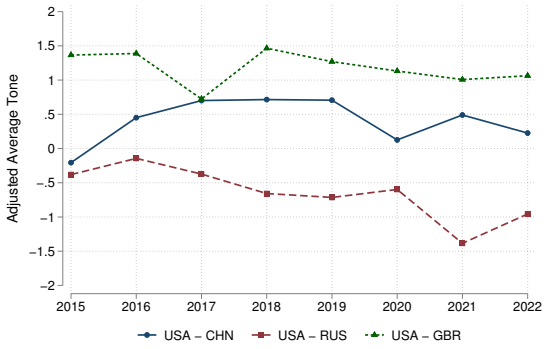
(b) Tone Distribution



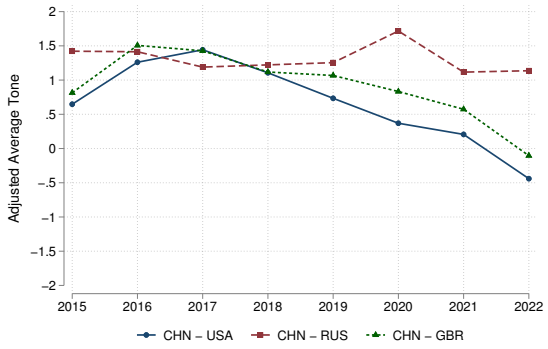
(c) Intertemporal Pattern of the Tone Measure

Figure 4: Distribution of Tone Measures

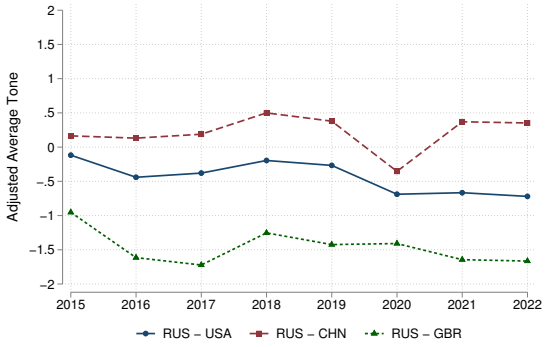
Notes: Panel (a) displays the distribution of positive and negative tone values. Panel (b) presents the distribution of the dependent variable—the monthly average tone of media outlets. Panel (c) illustrates the temporal pattern of the tone measure, with the 95% confidence interval shaded in gray.



(a) Source Country: USA



(b) Source Country: CHN



(c) Source Country: RUS

Figure 5: Tone Measures over Time by Country

Notes: Panel (a) displays the average tone of news articles issued by the United States covering issues in China, Russia, and the United Kingdom. Panel (b) presents the average tone of news articles issued by China covering issues in the United States, Russia, and the United Kingdom. Panel (c) shows the average tone of news articles issued by Russia covering issues in the United States, China, and the United Kingdom. In all panels, tone is adjusted by removing theme-by-year-month fixed effects estimated over the full dataset.

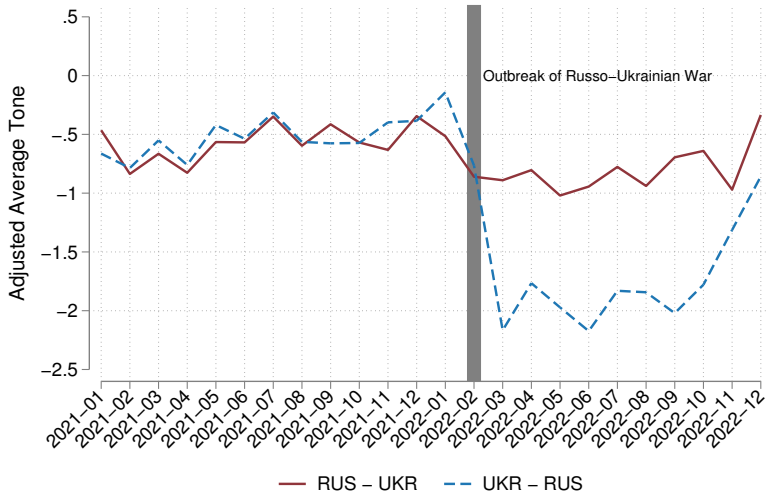
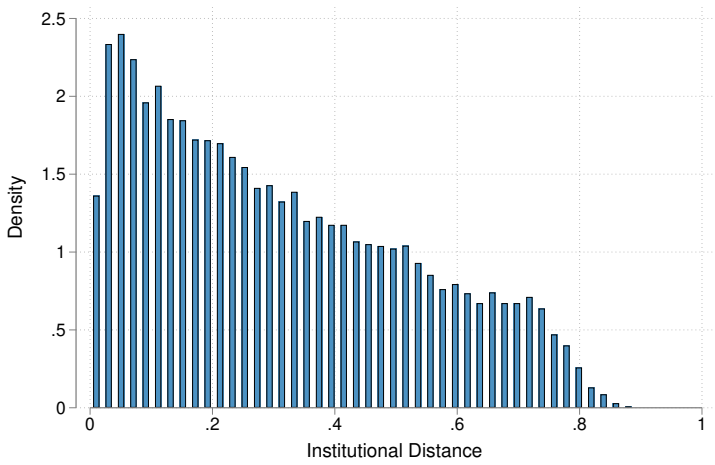
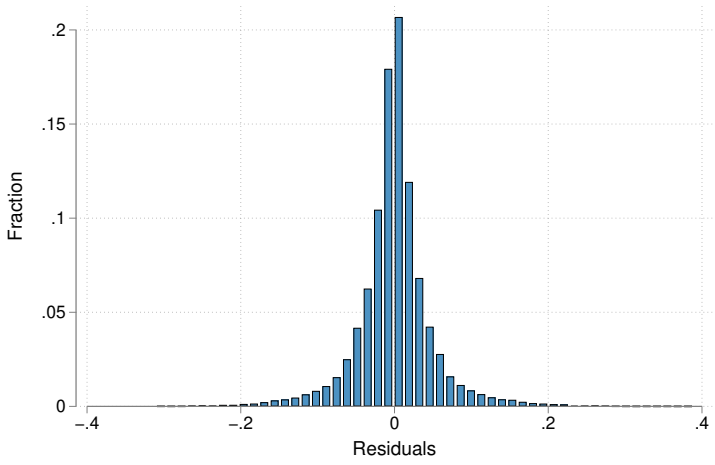


Figure 6: Changes in Media Tone Between Russia and Ukraine

Notes: The figure shows the changes in adjusted tones of Russian and Ukrainian media when they reported on each other. The red solid line represents the sentiment of the Russian media toward Ukraine, while the blue dashed line indicates the sentiment of the Ukrainian media toward Russia. Tone is adjusted by removing theme-by-year-month fixed effects estimated over the full dataset.



(a) Distribution of Institutional Distance



(b) Variation in Institutional Distance Across Years

Figure 7: 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. Panel (b) displays the distribution of institutional distance residualized by country-pair fixed effects.

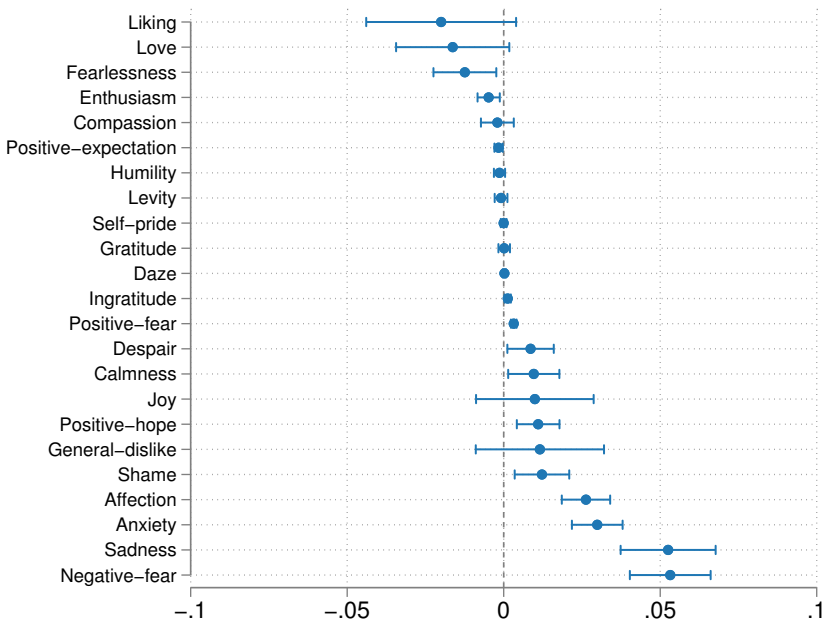


Figure 8: Sentiment Measures and Institutional Distance

Notes: This figure presents the estimated relationship between institutional distance and various sentiment measures, as classified by the WordNet-Affect lexicon. The coefficients and 95% confidence intervals from Table A3 are displayed.

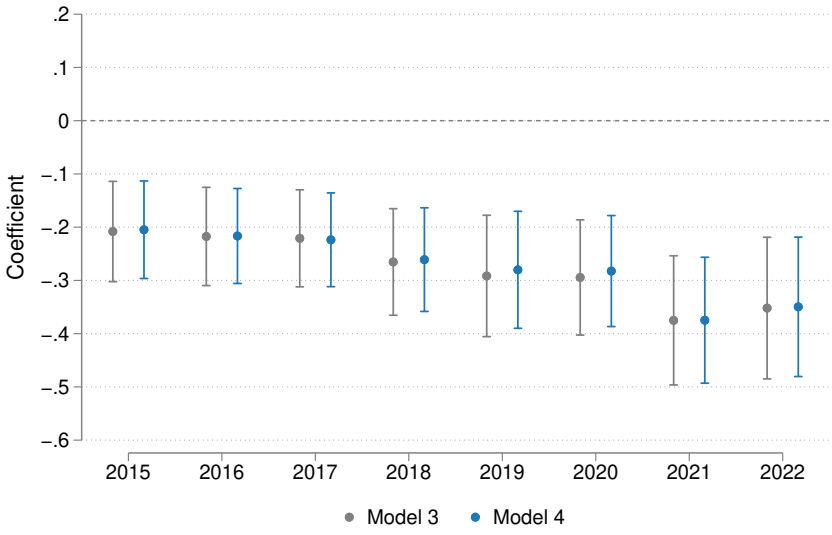


Figure 9: Year-by-Year Coefficients

Notes: This figure presents the year-by-year coefficients of tone on institutional distance based on the estimates from models (3) and (4) of Table A4. The point estimates and 95% confidence intervals are displayed.

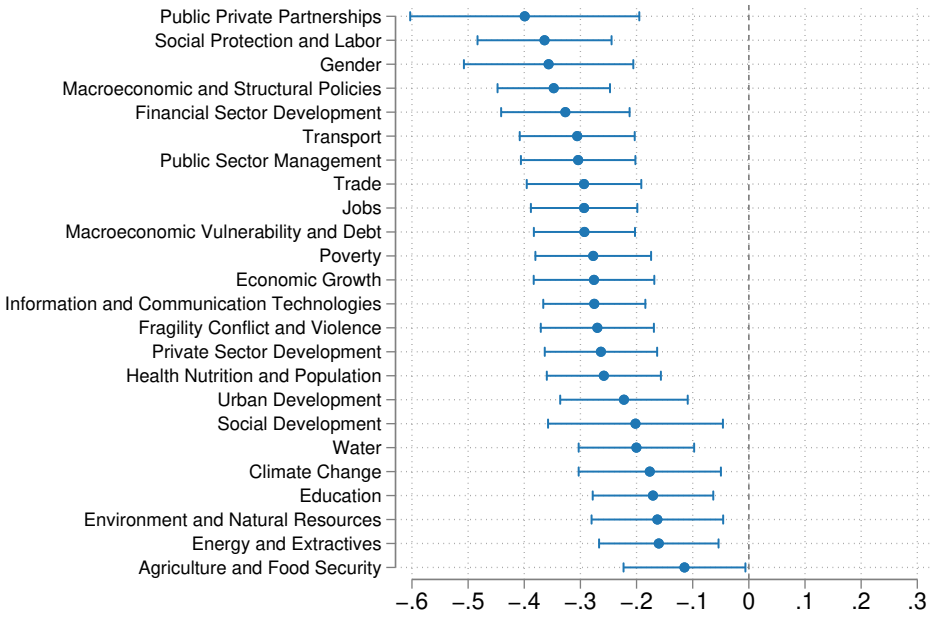
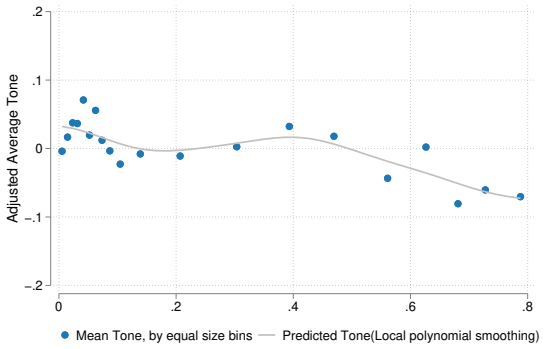
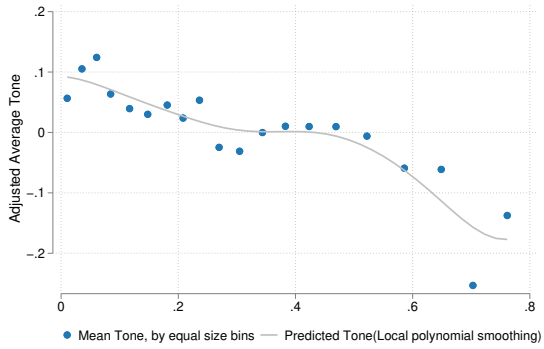


Figure 10: Heterogeneous Effects by Media Theme

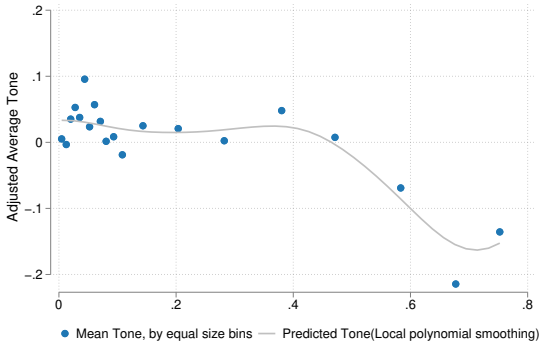
Notes: This figure shows the association between institutional distance and media sentiment across various themes, based on Level 1 of the World Bank's theme classification. The coefficient estimates and 95% confidence intervals from Table A16 are displayed.



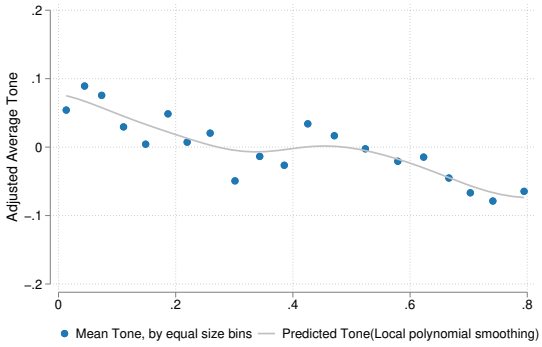
(a) Liberal Democracies as Source Countries



(b) Non-Liberal-Democracies as Source Countries



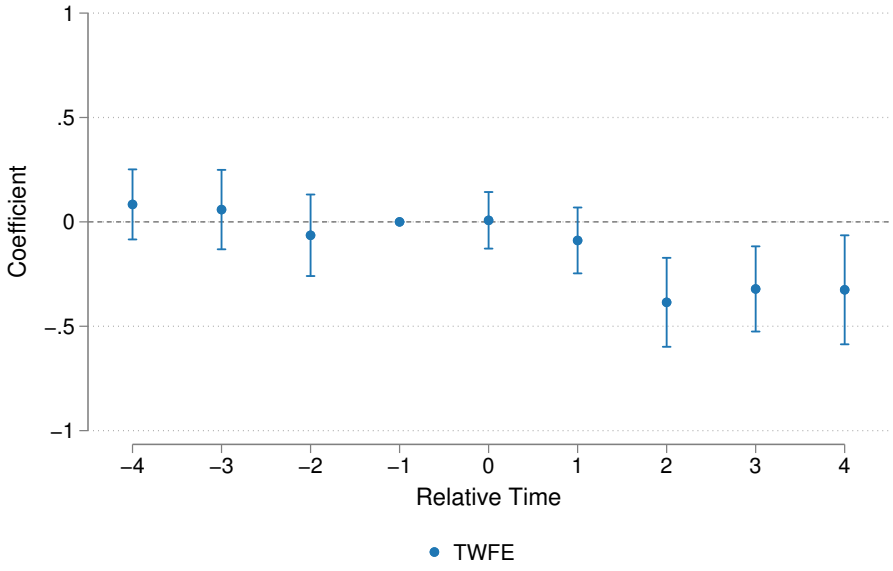
(c) Liberal Democracies as Target Countries



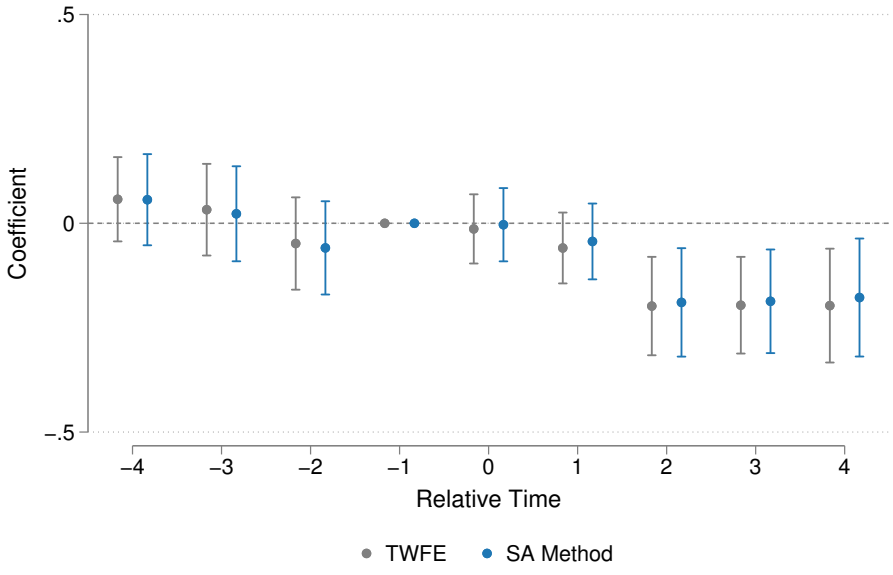
(d) Non-Liberal-Democracies as Target Countries

Figure 11: Democracy and Media Bias

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$^{\text{Source Country}}_{c,t}$ and Liberal Democracy Indicator$^{\text{Target Country}}_{c,t}$. We group the observations into 20 bins and plot the mean of the adjusted tone against the institutional distance for four separate subsamples.



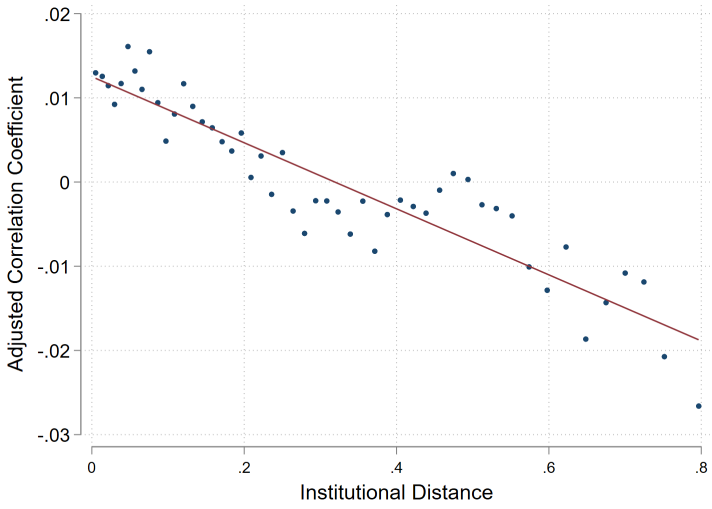
(a) Continuous Measure of Target's Democracy Level



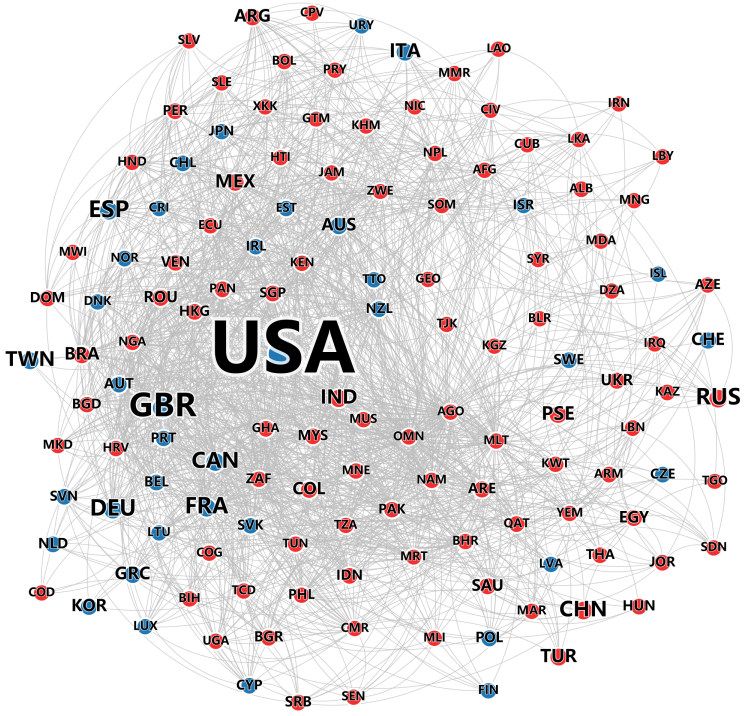
(b) Indicator for Whether Target is a Liberal Democracy

Figure 12: Staggered DID Event Study

Notes: This figure illustrates the dynamic effects identified through natural experiments triggered by regime changes. The coefficient estimates and 95% confidence intervals from models in Table A19 are displayed.



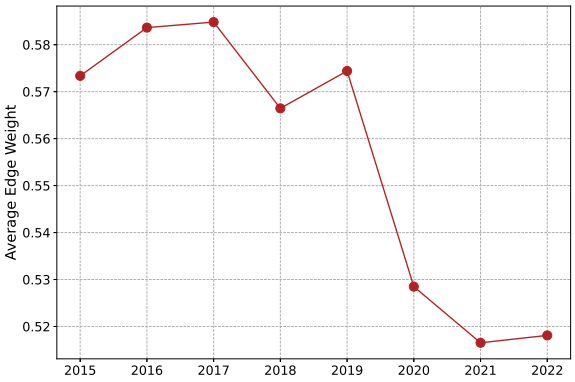
(a) Institutional Distance and Alliances



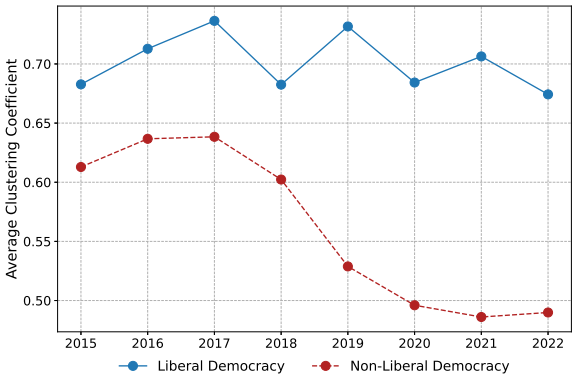
(b) Alliance Networks

Figure 13: Institutional Distance and Alliance Networks

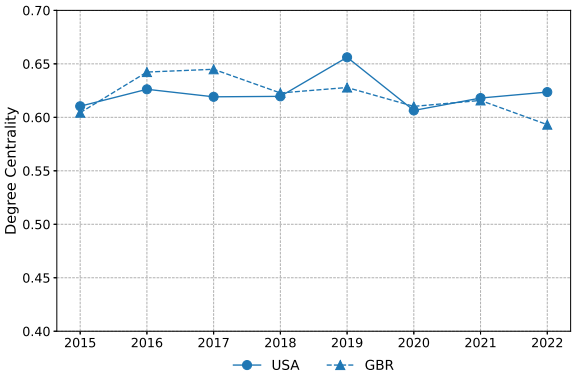
Notes: This figure illustrates how institutional distance shapes cross-country alignment in media coverage. Panel (a) presents a binned scatter plot of correlation coefficients (ρ) in media tone between pairs of countries, based on their coverage of other countries, plotted against bilateral institutional distance. The correlation coefficients (ρ) are adjusted by removing the target country, source country, and year fixed effects. Panel (b) presents a network graph in which nodes represent countries. Two countries are connected if the average correlation coefficient between 2015 and 2022 falls in the top 20th percentile. Blue nodes represent liberal democracies; red nodes denote non-liberal democracies. Node size reflects the volume of news articles issued by media outlets in that country.



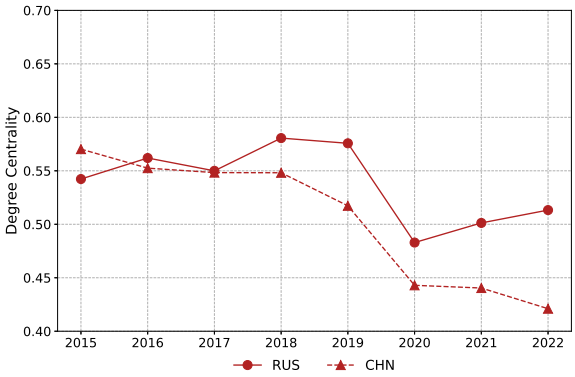
(a) Inter-Regime Connectivity



(b) Clustering Coefficient by Regime Type



(c) Degree Centrality: USA and GBR



(d) Degree Centrality: CHN and RUS

Figure 14: 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 non-liberal 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 non-liberal 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 regime networks.

Online Appendix

Table of Contents

A Additional Tables and Figures **66**

B Data Construction **87**

C Examples of News Articles from GDELT **97**

D Major News Producers by Country **99**

E Variable Definitions **100**

A Additional Tables and Figures

Table A1: Additional Summary Statistics

Panel A: Summary Statistics by Selected Country Pairs

	Tone				
Source Country - Target Country	Mean	Median	Std	P25	P75
USA - GBR	-.800	-.633	3.245	-2.710	1.297
GBR - USA	-.896	-.858	3.159	-3.107	1.331
USA - CHN	-1.551	-1.408	3.235	-3.548	.571
CHN - USA	-1.379	-1.237	3.019	-3.090	.398
USA - RUS	-2.546	-2.400	3.387	-4.706	-.284
RUS - USA	-1.912	-1.887	2.840	-3.571	-.183
CHN - RUS	-.424	-.263	3.564	-2.463	1.823
RUS - CHN	-1.467	-1.306	3.022	-3.246	.469
USA - IRN	-3.079	-2.825	3.660	-5.405	-.502
IRN - USA	-2.833	-2.859	3.019	-4.782	-.879

Panel B: Summary Statistics by Top 10 Themes

	Tone				
Theme	Mean	Median	Std	P25	P75
Public Sector Management	-2.579	-2.473	3.785	-4.952	.000
Health Nutrition and Population	-2.470	-2.290	3.785	-4.764	.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	.771
Education	-1.405	-1.205	3.681	-3.596	.991
Private Sector Development	-1.624	-1.401	3.622	-3.791	.775
Transport	-2.037	-1.812	3.754	-4.444	.575
Jobs	-.740	-.563	3.377	-2.722	1.440
Water	-1.652	-1.393	3.600	-3.846	.720
Financial Sector Development	-1.633	-1.362	3.377	-3.554	.544

Table A2: 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
Adjusted R-squared	0.181	0.194	0.293	0.306
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 the percentage of positive and negative words, respectively, in news articles. We aggregate tone measures at the outlet–target country–theme–month level. Other model specifications follow the convention in Table 2. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A3: Emotional Sentiments (WNA)

	Dependent Variable				
	Joy	Love	Affection	Liking	Enthusiasm
Institutional Distance	0.010	-0.016*	0.026***	-0.020	-0.005***
	(0.010)	(0.009)	(0.004)	(0.012)	(0.002)
Observations	81,288,425	81,288,397	81,288,434	81,288,416	81,288,434
Adjusted R-squared	0.158	0.129	0.051	0.220	0.103
	Gratitude	Self-pride	Levity	Calmness	Fearlessness
Institutional Distance	0.000	-0.000	-0.001	0.010**	-0.012**
	(0.001)	(0.000)	(0.001)	(0.004)	(0.005)
Observations	81,288,434	81,288,434	81,288,434	81,288,434	81,288,434
Adjusted R-squared	0.089	0.046	0.117	0.108	0.102
	Positive-expectation	Positive-fear	Positive-hope	Negative-fear	Sadness
Institutional Distance	-0.002**	0.003***	0.011***	0.053***	0.053***
	(0.001)	(0.000)	(0.003)	(0.007)	(0.008)
Observations	81,288,434	81,288,434	81,288,434	81,288,427	81,288,426
Adjusted R-squared	0.065	0.026	0.090	0.045	0.089
	General-dislike	Ingratitude	Shame	Compassion	Humility
Institutional Distance	0.012	0.001***	0.012***	-0.002	-0.001
	(0.010)	(0.001)	(0.004)	(0.003)	(0.001)
Observations	81,288,424	81,288,434	81,288,434	81,288,434	81,288,434
Adjusted R-squared	0.119	0.082	0.080	0.106	0.119
	Despair	Anxiety	Daze		
Institutional Distance	0.009**	0.030***	0.000		
	(0.004)	(0.004)	(0.000)		
Observations	81,288,434	81,288,434	81,288,434		
Adjusted R-squared	0.074	0.062	0.138		
Target Country × Theme × Year FEs	Yes	Yes	Yes		
Outlet × Theme × Year FEs	Yes	Yes	Yes		

Notes: This figure presents the relationship between institutional distance and a wide range of sentiment measures classified by the WordNet-Affect lexicon. Sentiments are measured in each news article and we aggregate the sample to the media outlet, target country, theme and month level. Standard errors in parentheses are clustered at the country pair level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A4: Dynamic Effects

	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
Adjusted R-squared	0.278	0.291	0.278	0.291
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 2.

Table A5: 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
Adjusted R-squared	0.300	0.304	0.330	0.339	0.343
Number of Country Pairs	14185	14172	14340	14120	14101
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 different fixed effects in addition to the baseline models in Table 2. Other variable construction and model specifications follow the convention in Table 2. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A6: Alternative SE Cluster 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	Source country	Source country × Year	Outlet	Outlet × Year
Observations	81,288,434	81,288,434	81,288,434	81,288,434	81,288,434
Adjusted R-squared	0.291	0.291	0.291	0.291	0.291
Number of Cluster 1	14,379	176	176	55,240	55,240
Number of Cluster 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 standard errors. Other variable construction and model specifications follow the convention in Table 2. *, **, 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
Adjusted R-squared	0.289	0.301	0.150	0.158
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), both derived from the GDELT GKG 2.0 database. Other variable construction and model specifications follow the convention in Table 2. *, **, 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 (PolityV Index)			-0.014***	-0.014***
			(0.003)	(0.003)
Observations	81,288,434	81,288,434	57,984,858	57,980,165
Adjusted R-squared	0.278	0.291	0.280	0.294
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 estimation by using different measures of institutional distance. Other variable construction and model specifications follow the convention in Table 2.

Table A9: Robustness Checks on Translation-Related Measurement

	Dependent Variables			
	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
Adjusted R-squared	0.191	0.291	0.291	0.309
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 1 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 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 2.

Table A10: Domain Source Identification 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 Rule	DNS	ChatGPT	Wiki	Top Level Domain
Observations	81,198,262	78,975,513	81,202,902	1,693,339	35,135,302
Adjusted R-squared	0.291	0.292	0.291	0.228	0.250
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 by using different domain source identification strategies. Other variable construction and model specifications follow the convention in Table 2.

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
Adjusted R-squared	0.295	0.297	0.303
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 presents robustness checks to the baseline estimation by 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 2.

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
Adjusted R-squared	0.260	0.289	0.259
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	No
Outlet × Theme × Year FEs	Yes	Yes	Yes

Notes: This table presents robustness checks to the baseline estimation by removing observations related to the United States. Other variable construction and model specifications follow the convention in Table 2. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A13: 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
Adjusted R-squared	0.304	0.314	0.331
Number of Country Pairs	14,200	14,200	14,200
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: We replicate the baseline results in Table 2 using data from the GDELT Event 2.0 Database. The sample is at the media outlet, target country, theme (measured with event types), and month level. 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 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
Adjusted R-squared	0.208	0.219
Number of Country Pairs	14379	14379
Target Country × Theme × Year FEs	Yes	Yes
Outlet × Year FEs	Yes	No
Outlet × Theme × Year FEs	No	Yes

Notes: This table presents robustness checks to the baseline estimation by collapsing the baseline sample to country-pair-by-theme-by year level. Other variable construction and model specifications follow the convention in Table 2. *, **, 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
Adjusted R-squared	0.293	0.293	0.291
Target Country × Theme × Year FEs	Yes	Yes	Yes
Outlet × Theme × Year FEs	Yes	Yes	Yes

Notes: This table presents robustness checks to the baseline estimation by including control variables related to trade relations. Other variable construction and model specifications follow the convention in Table 2. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table A16: Media Bias by Theme

	Dependent Variable: Monthly Average Tone	
	(1)	(2)
Institutional Distance × Information and Communication Technologies	-0.285***	-0.275***
	(0.0495)	(0.0464)
Institutional Distance × Social Development	-0.153*	-0.202**
	(0.0877)	(0.0794)
Institutional Distance × Transport	-0.310***	-0.306***
	(0.0516)	(0.0522)
Institutional Distance × Water	-0.232***	-0.200***
	(0.0543)	(0.0525)
Institutional Distance × Agriculture and Food Security	-0.157***	-0.115**
	(0.0592)	(0.0554)
Institutional Distance × Macroeconomic and Structural Policies	-0.246***	-0.347***
	(0.0668)	(0.0511)
Institutional Distance × Education	-0.152***	-0.171***
	(0.0541)	(0.0548)
Institutional Distance × Economic Growth	-0.160**	-0.276***
	(0.0779)	(0.0548)
Institutional Distance × Energy and Extractives	-0.0766	-0.160***
	(0.0823)	(0.0543)
Institutional Distance × Environment and Natural Resources	-0.261***	-0.163***
	(0.0608)	(0.0598)
Institutional Distance × Climate Change	-0.208***	-0.176***
	(0.0786)	(0.0646)
Institutional Distance × Gender	-0.168	-0.357***
	(0.109)	(0.0770)
Institutional Distance × Health Nutrition and Population	-0.338***	-0.258***
	(0.0576)	(0.0518)
Institutional Distance × Poverty	-0.237***	-0.277***
	(0.0627)	(0.0526)
Institutional Distance × Public Sector Management	-0.409***	-0.304***
	(0.0575)	(0.0520)
Institutional Distance × Social Protection and Labor	-0.358***	-0.364***
	(0.0672)	(0.0609)
Institutional Distance × Trade	-0.247***	-0.294***
	(0.0635)	(0.0520)
Institutional Distance × Urban Development	-0.355***	-0.222***
	(0.0993)	(0.0579)
Institutional Distance × Macroeconomic Vulnerability and Debt	-0.191***	-0.293***
	(0.0676)	(0.0459)
Institutional Distance × Financial Sector Development	-0.379***	-0.327***
	(0.0917)	(0.0584)
Institutional Distance × Private Sector Development	-0.205***	-0.263***
	(0.0558)	(0.0511)
Institutional Distance × Public Private Partnerships	-0.319***	-0.399***
	(0.119)	(0.104)
Institutional Distance × Fragility Conflict and Violence	-0.257***	-0.270***
	(0.0585)	(0.0514)
Institutional Distance × Jobs	-0.192***	-0.293***
	(0.0560)	(0.0483)
Observations	81288434	81288434
Adjusted R-squared	0.278	0.291
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 media bias by news article theme. We consider first level of World Bank Theme Category. Other variable construction and model specifications follow the convention in Table 2.*, **, 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 timing of the democratic erosion in the natural experiments as shown in Table A18. We show the treatment countries in our sample period between 2015 and 2022.

Table A18: Autocratization and Media Bias

	Dependent Variable: Monthly Average Tone	
	(1)	(2)
Autocratization × Liberal Democracy Index (Target Country)	-0.127**	
	(0.060)	
Autocratization × Liberal Democracy Indicator (Target Country)		-0.089***
		(0.032)
Observations	60,330,253	60,330,253
Adjusted R-squared	0.372	0.372
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 result of democratic erosion with a staggered difference-in-difference design. At this step, we drop the always-treated group and close autocracy countries that 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: Dynamic Effect of Democratic Erosion

	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
Adjusted R-squared	0.372	0.372	0.372
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 result of democratic erosion with a staggered difference-in-difference design. At this step, we drop the always-treated group and close autocracy countries that 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 A20: State Media and Media Bias

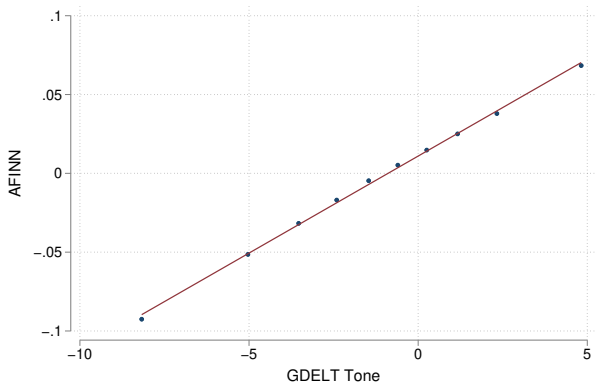
	Dependent Variable: Monthly Average Tone	
	(1)	(2)
Institutional Distance	-0.231***	-0.228***
	(0.043)	(0.041)
Institutional Distance × State Media Indicator	-0.295***	-0.293***
	(0.079)	(0.079)
Observations	81,288,434	81,288,434
Adjusted R-squared	0.278	0.291
Number of Country Pairs	14,379	14,379
Target Country × Theme × Year FEs	Yes	Yes
Outlet × Year FEs	Yes	No
Outlet × Theme × Year FEs	No	Yes

Notes: This table presents the heterogeneous correlation between institutional distance and media sentiment, conditional on media ownership. Ownership of each media outlet is labeled by the GPT 4o model. Other variable construction and model specifications follow the convention in Table 2. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

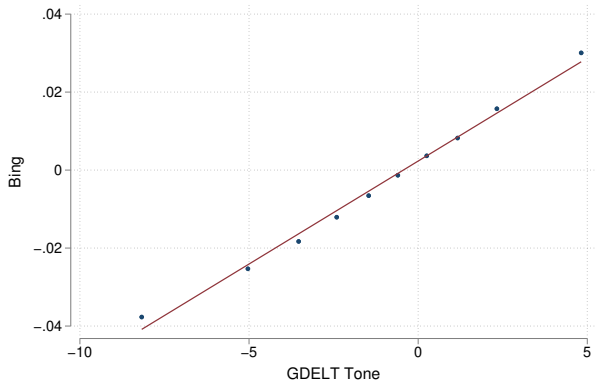
Table A21: Media Alliance and Diplomatic and Economic Alignment

	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
Adjusted R-squared	0.319	0.301	0.336
Source Country FEs	Yes	Yes	Yes
Target Country FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes

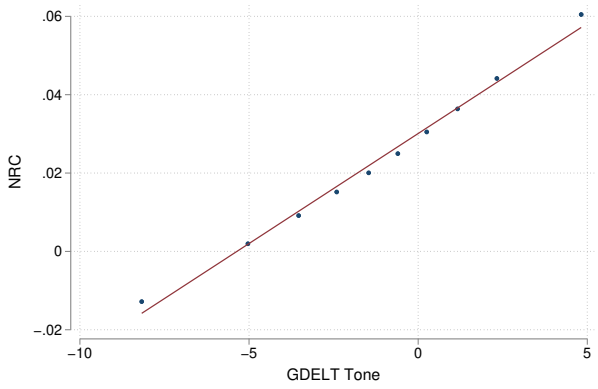
Notes: This table reports results from regressions estimating the predictive relationship between the media alliance measure and subsequent bilateral alignment outcomes. The independent variable is the media alliance index between country i and country j in year $t-1$, constructed from the correlation 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). All other model specifications follow those reported in Table 2. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.



(a) AFINN



(b) Bing Liu and collaborators



(c) NRC

Figure A1: Validation of the Tone Measures

Notes: This figure shows the validation of GDELT sentiment variables, with a random sample comprising of 157,779 new article texts. We construct three sentiment variables based on commonly used lexicon dictionaries and plot the correlation between each of those three variables against the GDELT sentiment variable. We show the scatter plot in ten bins. Panels (a), (b), and (c) show the sentiment variable based on AFINN lexicon, lexicon from Bing Liu and collaborators, and NRC lexicon, respectively. The Pearson correlation coefficients range between 0.66 and 0.81.

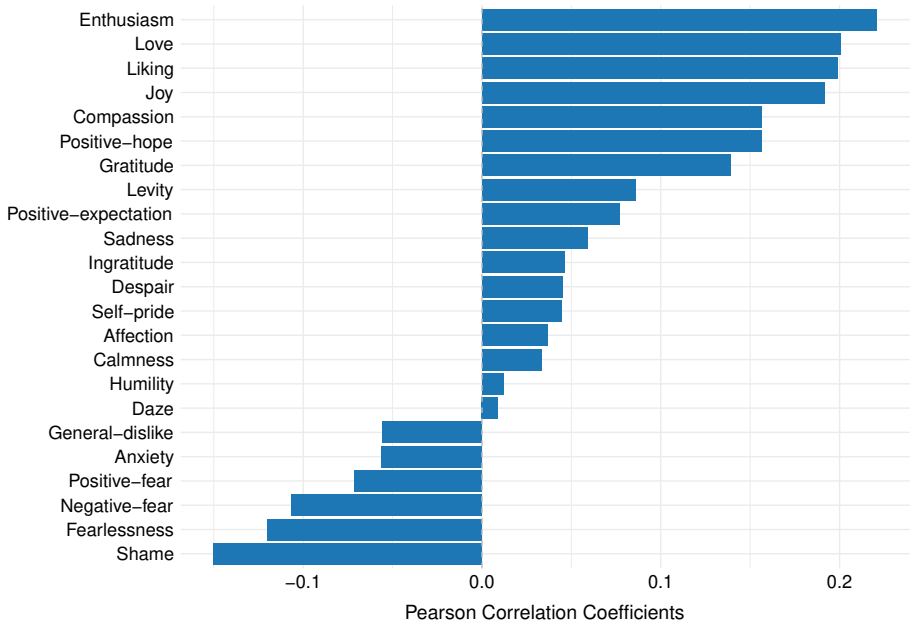


Figure A2: Correlations between Tone Measures and WNA Sentiment Measures

Notes: This figure presents Pearson correlations between our tone measures from the main analysis and WordNet-Affect (WNA) sentiment measures. Using a sample comprising of random selected 50,000 articles per month from the GDELT GKG 2.0 datavase, we compute the proportion of words associated with each WNA sentiment per article and analyze their correlation with tone measures. The results reveal strong positive correlations between positive sentiments — such as "enthusiasm," "love," "liking," and "joy" — and tone measures, while demonstrating strong negative correlations with negative sentiments, including "shame," "fearlessness," "positive-fear," "negative fear," and "anxiety." These consistent patterns lend credibility to our use of WNA sentiment measures as nuanced indicators of sentiment in the news text.

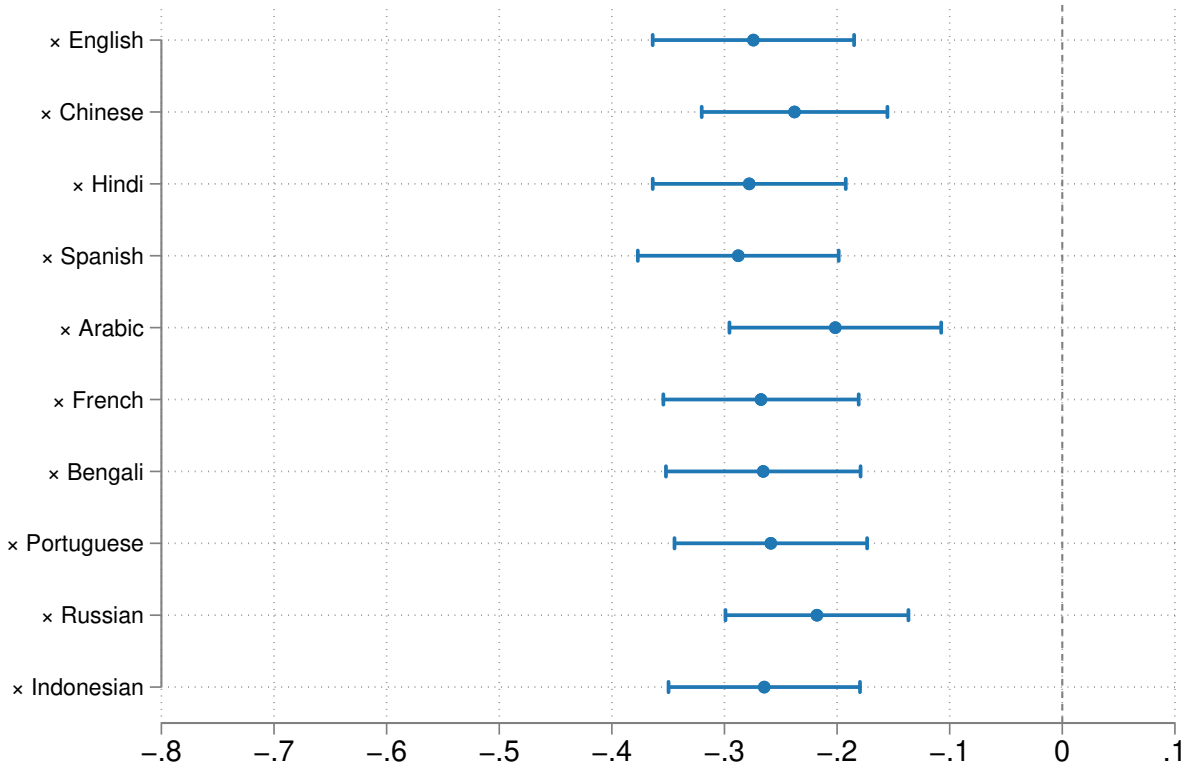


Figure A3: Robustness Check: Dropping One Major Language at a Time

Notes: This figure reports the results from excluding one major language at a time, to ensure that the findings are not driven by any single language. The coefficients and 95% confidence intervals are displayed.

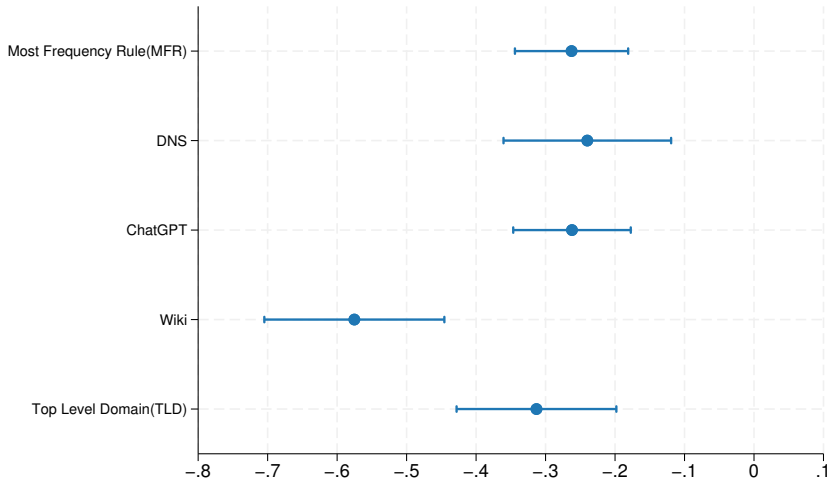
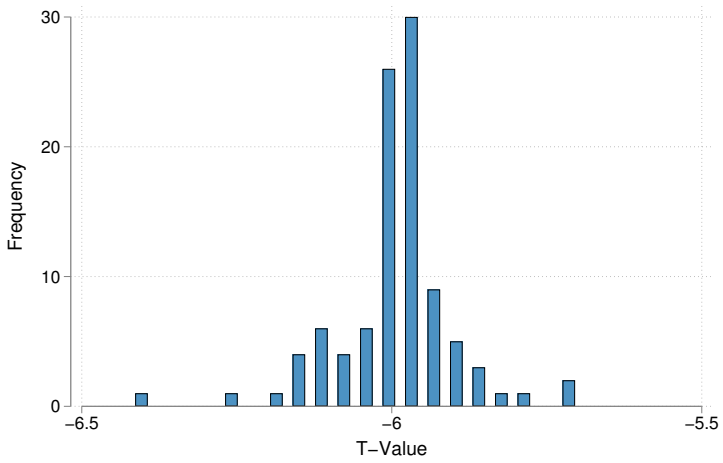
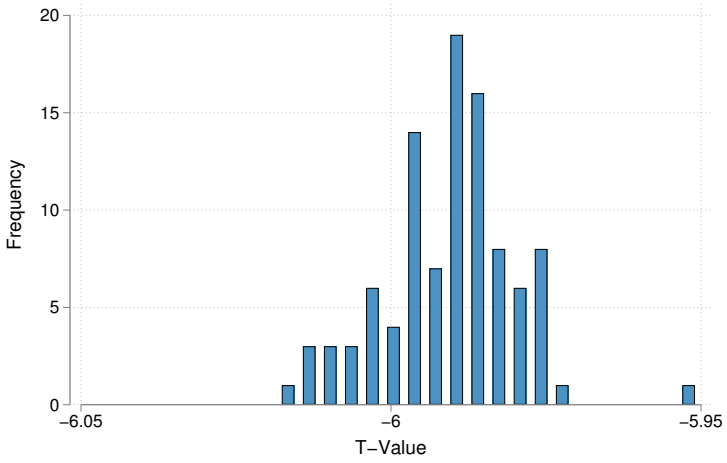


Figure A4: Results Using Different Domain Source Identification Methods

Notes: This figure presents the baseline results using different media source country identification rules. The coefficients and 95% confidence intervals are displayed. More details of the estimation can be found in Online Appendix Table A10.



(a) Excluding a Top 100 Country Pair once a Time



(b) Randomly Excluding 5% Sample Once a Time

Figure A5: Robustness Checks Based on Alternative Samples

Notes: This figure shows the robustness checks based on alternative samples. Panel (a) shows the t-values for the coefficient of interest in the sample excluding one of top 100 country-pair in terms of observations. Panel (b) shows results based on sample randomly excluding 5% of the observations once a time. We repeat this practice 100 times and plot the t-values for the coefficient of interest in each sample.

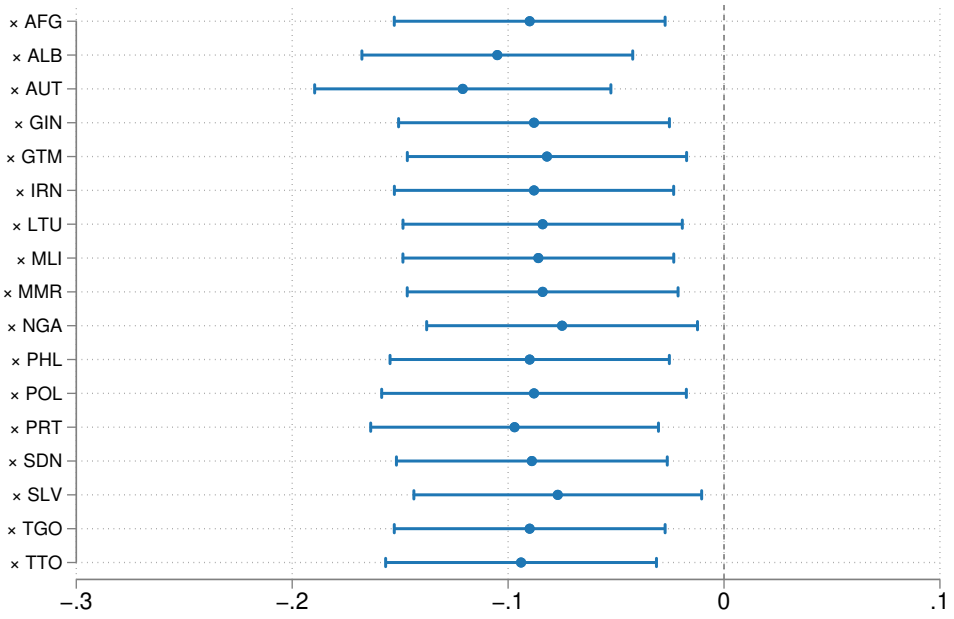


Figure A6: Robustness Check for DID Estimation

Notes: This figure presents robustness checks for the DID estimation results by excluding one democratic erosion case at a time. The coefficient estimates and 95% confidence intervals are displayed.

B Data Construction

B.1 Construction of 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 which should be highlighted here. Firstly, the Event 2.0 dataset only includes news articles where actors and actions are clearly identified. As of 2022, the Mentions table in the GDELT Event 2.0 dataset contains around 667 million events extracted from 420 million news articles. In contrast, the GKG 2.0 collects more than 1.4 billion news articles from more than 15,000 media outlets, providing much more detailed information. Secondly, the GKG 2.0 identifies every country mentioned in a news article, enabling us to focus on news related to a unique target country. Lastly, 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 replicate our baseline results using data constructed from the the GDELT Event 2.0 database. Results displayed in Table A13 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 database.

B.1.2 Data Collection and Cleaning

We obtained the raw data of GKG 2.0 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. Firstly, 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 by target country by theme by

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. News articles without any theme included in the World Bank Group Topical Taxonomy were removed in the aggregation process. At the aggregated level, we calculated the monthly average sentiment of news articles from each media outlet that targeted a given country with 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 A1. We also provide consistent results using the other two sentiment measures provided by the GKG 2.0 dataset in Table A7.

The aggregated data includes around 100 million observations at the media outlet-target country-theme-month level. Each observation represents the sentiment from a media outlet to 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. This step reduced our sample to around 80 million observations. We then further restricted the baseline sample to 81,288,434 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 1 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 if 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 included and 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.

Firstly, 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. Secondly, 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 official 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 in this study was the identification of source countries for tens of thousands of news outlets. This task was complicated by the difficulty of accessing consistent and 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. Additionally, 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 multi-step 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 (top-level domains, TLDs) to determine the source country. However, it is acknowledged that many media outlets frequently opt for generic suffixes (.com, .net, .org) rather than country-specific domains. Therefore, this method alone was insufficient for identifying source countries in today's digital media landscape.

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" (MFR) 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 data, the British newspaper *Daily Mail* (www.dailymail.co.uk) frequently covers the United States over domestic UK issues.

To ensure greater accuracy, we relied on the ChatGPT-4o API as our primary method 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. Additionally, as a robustness check, we used the most frequently reported country for each outlet as an alternative source identification measure.

In Figure A4 and Table A10, we show the baseline results using different source country identification methods. The estimated effects are comparable across different methods.

B.2 Construction of 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 2015–2022. The initial data collection utilized the Herre Database on Political Leaders as 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 ensure accuracy, we specifically identified leaders who wielded actual executive authority rather than ceremonial roles, taking into account various political systems including presidential, parliamentary, and one-party states. For instance, in parliamentary monarchies like the United Kingdom, we listed the Prime Minister rather than the monarch, while in one-party states like 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 for articles featuring each leader on Google News and compiled the names (and nicknames) most frequently used by news media. Our focus was on documenting both standardized naming conventions and region-specific variations in leader identification, including formal titles, common abbreviations, and culturally specific designations.

Using this comprehensive name reference dictionary, we then identified relevant English media coverage in the GDELT dataset through a two-step verification process. Articles were considered relevant to a leader if: (1) the article headline must contain at least one documented name reference and (2) 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.

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 10 most frequently reported national leaders in our sample.

B.3 Construction of Event Sample

As a parallel project of the GDELT, 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 but not limited to actors, event types, event locations, and news sentiment. Detailed information on Event 2.0 data is available in the codebook provided by the GDELT official (https://blog.gdeltproject.org/gdelt-2-0-our-global-world-in-realtime/).

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

Firstly, we identify 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 extract the unique address values from these location attributes to designate the target country. Consequently, events with multiple locations are excluded from the final analysis. Additionally, we remove news articles that mention multiple target countries to reduce confounding factors.

Next, we map the source country for each media outlet using the same method employed in our main results. Finally, we aggregate the data at the Media Outlet $\times$ Target Country $\times$ Event Type $\times$ Year-Month level and calculate the average tone of the news articles. It is important to note that 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 apply the top level of this classification scheme, which includes 20 event types, to our data analysis. This aggregated data is utilized for the robustness check, with results presented in Table A13.

B.4 Sample Construction for Equity Market Analysis

We construct the sample for analyzing the responses of the equity market to media bias across countries. For our research, we focus exclusively on U.S. firms because measuring firm-level exposure to media bias requires information on a firm's revenue from other countries, and the data quality for foreign firms is generally low. We source stock price information from the Center for Research in Security Prices (CRSP) and construct a panel of U.S. firms at the firm-month level for the period between 2015 and 2022. We source firm headquarter location information from Compustat. This original sample consists of 464,885 firm-month observations with 7,898 distinct listed firms headquartered in the U.S.[17] Asset pricing factors, risk free rates, and market returns are from Kenneth R. French's website.

To gauge firm exposure to media bias from a specific country, we collect revenue information from the FactSet Revere database, which includes details on the global revenue distribution of listed firms. This information is gathered by FactSet Revere from corporate filings and is broken down into geographic country or region categories.[18] More details are provided in Huang et al. (2023). We collect the revenue share for each listed U.S. firm at the fiscal year level and calculate *Tone Exposure* following Equation 19. Observations with missing or zero *Tone Exposure* are due to missing revenue data from a given country and are automatically excluded. *Tone Source Country* follows Equation 20. The final sample includes 169,806 firm-month observations with 2,734 distinct listed U.S. firms between 2015 and 2022, complete with essential stock return information and measures of media exposure. Summary statistics of the sample are provided in the table below to support the findings presented in Table 10.

	N	Mean	SD	Median
Tone Exposure	169,806	-0.405	0.353	-0.322
Tone Source Country	169,806	-0.330	0.348	-0.246
Return	169,806	0.005	0.138	0.002
Return (CAPM)	169,806	-0.005	0.126	-0.006
Return (FF3)	168,489	-0.004	0.127	-0.005

[17] This dataset was retrieved from CRSP on December 23, 2023.
[18] This dataset was retrieved from FactSet Revere on December 23, 2023.

B.5 Sample Construction for Public Opinion Responses

We collect public opinion data from the Gallup World Poll Project, which is one of the most influential international surveys. Starting in 2006, Gallup's World Poll conducted ongoing surveys of people in over 160 countries and areas, representing more than 98% of the global adult population. This is achieved with randomly selected, 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 that are translated into the primary languages of each respective country.

The Gallup World Poll conducts surveys to identify the most urgent challenges confronting humanity and to monitor advancements in human development, including but not limited to questions related to Business and Economics, Government and Politics, and Public Health. These indicators assist researchers in understanding the broad context of national interests and establish specific relationships between indexes and lagging economic outcomes.

The Gallup World Poll Project asks respondents to evaluate their approval of leaders in different countries with binary answers ranging from Disagree to Agree. We removed respondents who refused to answer the question or those who answered Don't Know. This variable facilitates us in exploring the relationship between cross-country media bias and the sentiments of local people. To avoid the reverse causality problem, we use the approval variable one year after the media reports for each observation. In the analysis, we also include several control variables, including Gender, Income Level, Educational Level, and Employment Status.

The Gallup World Poll (2016-2023) used in our sample surveys people's approval of leadership in 9 target countries, including China, Germany, France, India, Iran, Japan, Russia, Turkey, and the United States of America. The respondents are from 147 countries and areas. In total, 779 country pairs are identified. Summary statistics of this sample are provided in the table below to support the results presented in Table 11.

	N	Mean / %	SD	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 collected by the GDELT project from CNN with a focus 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 collected by the GDELT project from China.org.cn, which highlighted the problem of gun violence in the USA. 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.

E Variable Definitions

Panel A: Variables at Domain - Target - 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)
Common Language	A dummy variable indicating if two countries share the same official language.	CEPII
Geographic Distance	Geographic distance between 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 the UN Comtrade. It is determined by summing the exports from country i to country j and the imports from country j to country i, and 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 the UN Comtrade. It is determined by subtracting the imports from country j to country i from the exports from country i to country j, and then dividing this difference by the GDP of country i.	UN Comtrade
Trade Sanction	A dummy variable indicating if country i was sanctioned by country j in the past year, with data from the Global Sanctions Data Base (GSDB).	GSDB

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 if 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 if a country experiencing democratic backlash in or before the current year.	V-Dem Project
Freedom of Traditional Media	An index (v2mecenefm) from the V-Dem project assessing does the government directly or indirectly attempted to censor the print or broadcast media in the past year? A higher value means a higher freedom of traditional media.	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 a higher freedom of the Internet.	V-Dem Project
State-owned Print Media	An index (v2medstateprint) from the V-Dem project assessing out of the top four national print media with the highest readership in th past year, how many are state-owned.	V-Dem Project
State-owned Broadcast Media	An index (v2medstatebroad) from the V-Dem project assessing out of the top four national broadcast media with the largest audience in th past year, how many are state-owned.	V-Dem Project
Monthly Growth in Nighttime Light Intensity	The measurement of growth in Monthly Light Intensity is determined by taking the average light intensity of the past month and dividing it by the average light intensity from two months ago.	VIIRS-DNB
High Inflation Indicator	High inflation indicator is equal to 1 when the inflation rate in 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 database.	World Bank
Domestic Turmoil Index	Share of domestic turmoil events in the past month, generated by approaches from Amarasinghe (2022) with data from GDELT Event 2.0.	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

Panel E: Variables in Equity Market Analysis

Variable	Definition	Source
Tone Exposure	The firm-month level measure. $Tone\ Exposure_{i,t} = \sum_{c=1}^{C} Tone_{c,t} \times Rev_{i,c,t}$. $Rev_{i,c,t}$ is a firm's revenue share from a given country c. $Tone_{c,t}$ is the average tone of the news articles from media outlets from country c towards United States in month t.	GDELT GKG 2.0, Factset Revere
Tone Source Country	The firm-month level measure. $Tone\ Source\ Country_{i,t} = \sum_{c=1}^{C} Tone_{c,t}^{Source} \times Rev_{i,c,t}$. $Tone_{c,t}^{Source}$ is the average tone of the media outside country c towards United States in month t.	GDELT GKG 2.0, Factset Revere
Return	The monthly stock returns for US listed firms.	CRSP
Return (CAPM)	The monthly abnormal stock returns for US listed firms adjusted by the market model.	CRSP, Kenneth R. French website
Return (FF3)	The monthly abnormal stock returns for US listed firms adjusted by the Fama-French 3 factor model.	CRSP, Kenneth R. French website