Beyond the Fundamentals: Cross-Country Media Narrative Dispersion and Global Capital Flows*

Isha Agarwal† Wentong Chen‡ Eswar Prasad§

March 3, 2026

Abstract

This paper provides the first empirical evidence that cross-country variation in domestic media narratives about a destination country influences global institutional investment flows. Using natural language processing methods, including transformer-based sentiment models and LLMs, applied to large-scale newspaper text from multiple countries, we measure narrative disagreement across investor countries about the same foreign economy. Drawing on more than one million newspaper articles from 39 outlets across 16 economies, we construct country-specific measures of media attention and sentiment toward China, an increasingly important investment destination with severe information frictions. We document large and persistent cross-country dispersion in sentiment, even when media cover the same topics. A counterfactual decomposition reveals that this dispersion is driven almost entirely by differences in within-topic sentiment rather than topic attention, indicating heterogeneous framing of common information. Linking narratives to behavior, we show that exogenous changes in domestic media sentiment significantly affect cross-border portfolio flows to China using a shift-share identification strategy. Our findings establish domestic media narratives as an important channel shaping belief formation and global capital allocation.

JEL Classification Numbers: F30, G11, G15
Keywords: media narratives, cross-border flows, institutional investors, portfolio investment in China, textual analysis, natural language processing

---

*The paper was previously circulated as "Beyond the Fundamentals: How Media-Driven Narratives Influence Cross-Border Capital Flows". The authors would like to thank Tony Cookson, Shimon Kogan (discussant), Song Ma, Varun Sharma (discussant), Chenzi Xu (discussant), and conference participants at the RCFS Winter Conference 2026, NBER Chinese Economy Working Group Meeting Spring 2025, the 9th Monash-Paris-Warwick-Zurich-CEPR Text-As-Data Workshop, the 2025 Milan Conference on Information, Contracts and Firms, and the 17th FIW Research Conference 'International Economics'. Yicheng Yang provided excellent research assistance.

†Sauder School of Business, University of British Columbia, E-mail: isha.agarwal@sauder.ubc.ca

‡Department of Economics, Santa Clara University, Santa Clara, CA, 95053, U.S.A, E-mail: wchen10@scu.edu

§Dyson School of Applied Economics and Management and Department of Economics, Cornell University, Ithaca, NY, 14853, U.S.A, Brookings Institution, and NBER, E-mail: eswar.prasad@cornell.edu

1 Introduction

Economic agents rely on narratives, stories that organize and interpret complex economic events, when making decisions under limited attention and information frictions (Shiller, 2017, 2020). News media play a central role in shaping these narratives by selecting, framing, and interpreting information through editorial "gatekeeping" decisions (Shoemaker and Vos, 2009; Nimark and Pitschner, 2019; Chahrour, Nimark and Pitschner, 2021). A growing literature shows that media attention and sentiment affect asset prices, trading activity, and investor behavior (e.g., Tetlock, 2007; Engelberg and Parsons, 2011; Manela and Moreira, 2017; Bybee et al., 2024; Calomiris and Mamaysky, 2019; Baker et al., 2021). These forces are likely to be particularly salient in cross-border settings, where investors face severe information and interpretation frictions (Van Nieuwerburgh and Veldkamp, 2009).

Compared with domestic investing, cross-border portfolio allocation requires processing large volumes of unfamiliar and often opaque information, complicated by language barriers, institutional differences, and policy uncertainty. As a result, institutional investors may be especially reliant on intermediaries, such as domestic news media, to filter and interpret developments abroad. Despite the importance of media narratives in shaping investment dynamics, existing evidence remains overwhelmingly domestic in scope, focusing on media coverage of domestic assets or macroeconomic conditions, or examining sentiment toward different countries without studying heterogeneity in media narratives about the same country.[1] Consequently, we still lack evidence on whether media narratives about the *same foreign economy* differ systematically *across investor countries* and

[1] A growing literature uses textual analysis of news to construct sentiment, uncertainty, and risk indices in an international setting. Early work by Calomiris and Mamaysky (2019) demonstrates how news content and context extracted from global media sources can be used to explain asset price movements. Related studies construct country-level indices of economic uncertainty or risk using domestic newspapers, including the Economic Policy Uncertainty index of Baker et al. (2016), the World Uncertainty Index of Ahir et al. (2022), and the Geopolitical Risk index of Caldara and Iacoviello (2022). While these studies exploit international news coverage, they do not examine sentiment disagreement across countries about a common destination.

whether such differences matter for cross-border capital allocation.

To close this gap, we use modern natural language processing techniques, including transformer-based sentiment models and large language models, to construct cross-country indices of media narratives about China from a large panel of newspapers across investor countries. Drawing on more than one million articles, we measure both media attention and sentiment toward China as reflected in investors' home-country news coverage. This framework allows us to systematically compare how the same destination economy is portrayed across countries and to quantify persistent differences in narrative framing. We then link this narrative heterogeneity to international portfolio flows, showing that cross-country variation in media-based beliefs about a common foreign economy has economically meaningful effects on cross-border capital allocation.

We focus on China, an increasingly important destination for global institutional investors and a setting in which information frictions are especially pronounced. Investors face well-known challenges in accessing timely and reliable data on China, which are exacerbated by opacity in policymaking and periodic reductions in data dissemination.[2] Even when data are available, interpretation is costly due to language barriers, institutional complexity, and frequent regulatory changes. In this environment, domestic media narratives can play a central role in shaping investors' decisions by influencing both which aspects of China receive attention and how available information is interpreted.

Media-driven narratives influence investors' decisions at both the extensive and intensive margins. At the extensive margin, the volume of China-related news shapes the level of attention devoted to China by global investors. At the intensive margin, each article can affect investors by shaping their sentiment toward China. To capture these dimensions, we construct two country-level indices: an attention index and a sentiment index, using modern natural language processing methods applied to more than one million articles published in 39 English-language newspapers across 16 economies between 2007 and 2022.

[2]For example, *Bloomberg* reports that China stopped publishing daily global stock flows data in August 2024.

The attention index captures the share of news allocated to China while the sentiment index reflects the intensity of positive and negative sentiments in local newspapers about China. We believe this is the first paper to leverage such a vast and diverse collection of *global* newspaper data to construct media-based narrative indices about the *same* country.

We begin by documenting substantial cross-country dispersion in media sentiment about China. Some countries consistently portray China in a negative tone, while others adopt markedly more positive portrayals. This variation in aggregate sentiment can arise through two distinct channels. First, countries may differ in the topics they emphasize when reporting on China, with some topics being inherently more negative than others. Second, even conditional on similar topic coverage, media outlets may differ in how they frame and interpret the same topics. The richness of our dataset allows us to construct twelve topic-level share and sentiment indices across countries using an LLM-based classifier. This granular structure allows us to move beyond aggregate sentiment measures used in prior work and systematically decompose narrative disagreement into differences in topic attention and within-topic sentiment. In doing so, we provide new evidence on the sources of cross-country variation in international media narratives about China.

We document substantial cross-topic and cross-country variation in both what countries cover and how they frame those topics. Among China-related news, themes such as society and demographics, firms and corporate strategy, and geopolitics receive relatively high attention overall. Asian economies tend to devote significantly more coverage to these topics compared with other advanced economies. We also observe systematic differences in tone across topics. Coverage of human rights, health, and governance is markedly more negative on average than coverage of firms and corporate strategy or technology and industry policy. Moreover, advanced economies, including the U.S., the U.K., Canada, and Australia, tend to exhibit more negative sentiment across most topics relative to Asian economies.

We conduct a counterfactual decomposition exercise to assess the relative importance

of topic coverage and within-topic sentiment in explaining cross-sectional variation in aggregate media sentiment about China. For each investor country, we begin by expressing the aggregate sentiment index about China as a weighted average of within-topic sentiment, where topic shares serve as weights. We then construct two counterfactual sentiment indices: one that holds topic coverage fixed across investor countries while allowing within-topic sentiment to vary, and another that holds within-topic sentiment fixed while allowing topic coverage to vary. Comparing the cross-country variance of these counterfactual indices with that of the observed aggregate sentiment index allows us to quantify the contribution of each channel. We find that nearly all of the cross-country variation in aggregate sentiment is driven by differences in within-topic sentiment, suggesting that cross-country disagreement in media narratives about China reflects differences in how common information is processed rather than differences in topic coverage.

Do these differences stem from heterogeneous baseline priors about China, from systematic differences in how countries interpret common signals, or from differences in how strongly they update beliefs in response to new information? To address this question, we decompose each country's topic-level sentiment into components implied by a Bayesian updating framework. For each topic, we separate (i) belief-level heterogeneity, which includes persistent country-specific priors capturing inherited baseline stances and interpretation wedges reflecting systematic cross-country differences in mapping common public signals into sentiment, and (ii) updating heterogeneity, which captures cross-country variation in updating intensity, the relative weight placed on prior beliefs versus new information. The variance decomposition reveals substantial heterogeneity across topics in the relative importance of these channels. On average, belief-level heterogeneity (priors and interpretation combined) accounts for the majority of cross-country variation in sentiment, while updating heterogeneity becomes more important in topics characterized by more frequent or more measurable information flows.

Consistent with this decomposition, the cross-country belief level analysis shows

that persistent differences in narrative beliefs are systematically related to structural country characteristics. In particular, measures of economic interdependence, geopolitical positioning, institutional quality, and financial integration with China explain a substantial share of the variation in estimated country-specific belief levels. These findings reinforce the interpretation that baseline narrative lenses are not arbitrary, but instead reflect deeper economic, political, and financial linkages to China.

These findings raise a natural question: do cross-country differences in media narratives translate into differences in cross-border capital flows to China? To address this question, we combine our narrative measures with portfolio holdings data for open-ended mutual funds from Morningstar. Using quarterly fund-level data for approximately 20,000 funds domiciled in sixteen economies, we find that domestic media sentiment toward China has a statistically and economically significant effect on portfolio allocation decisions, even after controlling for China's macroeconomic and financial fundamentals. Holding constant these fundamentals, countries whose media portray China more negatively subsequently allocate less capital to Chinese assets, while more positive sentiment predicts stronger portfolio flows to China. The magnitude of these effects is economically meaningful: a one–standard deviation increase in the sentiment index is associated with a 0.96% increase in quarterly investment flows into China, corresponding to an annualized increase of 3.82%.

We confirm the robustness of this baseline relationship along several dimensions. First, to ensure that our results are not specific to transformer-based sentiment classification, we re-measure narrative tone using a traditional bag-of-words approach based on established financial dictionaries by exploiting the full text of articles. Second, to rule out the possibility that our results are driven by reliance on English-language media, we construct multilingual sentiment indices using language-specific FinBERT-style models applied to major domestic newspapers in Germany, France, Spain, Switzerland, and Austria. Third, to isolate narrative variation that is orthogonal to macroeconomic and financial funda-

mentals, we extract media narrative shocks from a structural VAR that conditions on key indicators of Chinese economic activity and financial conditions. These narrative shocks strongly predict subsequent cross-border capital flows, reinforcing the interpretation that media narratives affect investment decisions through framing and interpretation rather than through contemporaneous fundamentals alone.

To provide further evidence on the role of interpretation in shaping media narratives, we exploit the Arab Spring as a plausibly exogenous global shock to geopolitical narratives. The Arab Spring triggered a worldwide reassessment of political stability, governance, and regime durability, elevating the salience of geopolitical risk in media narratives well beyond the countries directly affected. Although China's economic fundamentals were unchanged, this shift in global political narratives led media outlets to reinterpret China through a more geopolitical risk lens. Consistent with this mechanism, we show that following the Arab Spring, countries whose media environments are more oriented toward geopolitical reporting experience a pronounced deterioration in sentiment toward China, despite no corresponding differential change in media attention to China. These narrative shifts are accompanied by relative declines in portfolio flows to Chinese assets, supporting a narrative-based transmission mechanism in which global shocks affect cross-border investment through country-specific framing and interpretation rather than changes in country fundamentals.

Complementary to the Arab Spring event study, we conduct an additional exercise to construct a more general narrative shock to China. The central identification challenge is to disentangle media driven sentiment shifts from changes that arise in response to China's macroeconomic or financial fundamentals. We construct an instrumental variable for domestic China sentiment using non China related coverage in the *Wall Street Journal* and interact these global narrative shocks with predetermined country level topic composition. This shift share design exploits exogenous variation in the global narrative environment while holding fixed countries' exposure shares. As a result, it isolates variation in investor

country media sentiment that arises from differential exposure to global narrative fluctuations rather than from contemporaneous China specific economic developments. We show that when media narratives about China moves for reasons unrelated to China's own economic or financial fundamentals, those movements still affect capital flows.

Finally, we study asymmetries in the effects of media narratives and the role of higher-order narrative measures. Consistent with a large literature on negativity bias (Holbrook, Krosnick, Visser, Gardner and Cacioppo, 2001; Soroka, 2006), we find that adverse media narratives about China exert a significantly stronger influence on institutional portfolio allocation than favorable narratives. This asymmetry suggests that downside information plays a dominant role in shaping cross-border investment decisions, especially in settings characterized by limited transparency. We also identify an independent role for media-implied risk. Increases in media-based risk perceptions about China, which capture heightened uncertainty rather than average sentiment, predict additional declines in investment flows even after controlling for macroeconomic and financial fundamentals. These results suggest that domestic media shape international capital allocation not only through average narrative tone, but also through second-moment perceptions of risk, reinforcing the importance of media narratives as a channel of belief formation in global finance.

Related Literature. Our paper contributes to the literature on news media, narratives, and financial market outcomes. While a large body of work shows that media affects financial market outcomes (e.g., Tetlock, 2007; Engelberg and Parsons, 2011; Manela and Moreira, 2017; Bybee et al., 2024; Calomiris and Mamaysky, 2019; Baker et al., 2021), this literature has remained overwhelmingly domestic in scope, focusing on media coverage of domestic assets or macroeconomic conditions.[3] Despite this extensive literature and the

[3]Existing research has also examined the impact of news media on investment dynamics and business cycles within closed economies (Gulen and Ion, 2016; Hassan, Hollander, Van Lent and Tahoun, 2019; Chahrour, Nimark and Pitschner, 2021; Flynn and Sastry, 2024; Bybee, Kelly, Manela and Xiu, 2024; Hu, 2024).

potential importance of media narratives in international settings, we still lack evidence on how media narratives about the *same foreign economy* differ across *multiple investor countries* and whether such differences matter for global capital allocation. We contribute to this literature by being the first to construct cross-country media-based narrative indices about the same destination economy using newspapers from multiple investor countries. We show that media narratives about the same economy are not uniform across countries, but instead differ systematically across investor countries and shape global capital flows.

Our paper also contributes to the growing text-as-data literature that uses advances in natural language processing to measure sentiment, risk, and uncertainty from large text datasets (Baker et al., 2016; Handley and Li, 2020; Ahir et al., 2022; Hassan et al., 2019, 2024; Arteaga-Garavito et al., 2024; Caldara and Iacoviello, 2022; van Binsbergen et al., 2024).[4] This literature has generated aggregate measures but it has largely focused on a single media environment. At the same time, a growing literature on belief dispersion shows that investors can disagree even when facing similar information, due to heterogeneous priors, attention, or interpretation (Cookson et al., 2020; Cookson and Niessner, 2020; Cookson et al., 2023, 2024). We bridge these literatures by introducing a media-based, cross-country measure of narrative disagreement about the same foreign economy and by decomposing this disagreement into differences in topic attention and within-topic sentiment. Methodologically, our analysis relies on modern natural language processing tools, including transformer-based sentiment models and LLMs, applied to large-scale newspaper text from multiple countries. Moreover, we develop a unified multilingual sentiment measurement framework based on language-specific FinBERT-style transformer

[4] Baker, Bloom and Davis (2016) develop economic uncertainty indices using counts of newspaper articles. Handley and Li (2020) create firm-level risk indices for U.S. firms using SEC filings. Ahir, Bloom and Furceri (2022) produce country-level uncertainty indices from Economist Intelligence Unit reports. Hassan, Hollander, Van Lent and Tahoun (2019) and Hassan, Schreger, Schwedeler and Tahoun (2024) use U.S. firms' earnings call transcripts and 10-K filings to measure political risks and perceived country risks. Arteaga-Garavito, Colacito, Croce and Yang (2024) construct climate attention indices from newspaper tweets, while Caldara and Iacoviello (2022) develop a geopolitical risk index using U.S. newspaper articles. van Binsbergen, Bryzgalova, Mukhopadhyay and Sharma (2024) exploits long historical newspaper archives to measure economic sentiment over very long time horizons. Hoberg and Manela (2025) review this literature.

models, enabling comparable narrative measurement across domestic media environments without relying on translation or English-only reporting. This integration allows us to study belief dispersion in an international setting and to link it directly to cross-border capital flows.

Finally, our paper contributes to the literature on determinants of cross-border capital flows.[5] We identify a previously overlooked determinant: local media narratives in investors' domicile countries. Existing studies emphasize information frictions as a key driver of international investment (Van Nieuwerburgh and Veldkamp, 2009; Andrade and Chhaochharia, 2010; Karolyi, Ng and Prasad, 2020), but typically treat these frictions as latent or proxy them with geographic, institutional, or technological distance. In contrast, we provide a direct, investor-facing measure of the information and interpretation environment, capturing not only the availability of information but also how common information is framed and interpreted in domestic media. By documenting that media narratives influence capital flows over and above traditional macro-financial drivers, our findings add a new dimension to the study of international investment behavior: belief formation shaped by domestic media narratives, rather than fundamentals alone.[6]

2 Measuring Media Narratives at the Micro Level

In this section, we describe our methodology for quantifying media narratives about China using natural language processing (NLP) techniques. We begin by outlining the data collection process and then detail the construction of our narrative indices. For each

[5] An extensive literature has studied the determinants of cross-border flows. Previous studies have highlighted various factors affecting these flows, including quality of governance (Leuz, Lins and Warnock, 2009), currency returns (Froot and Ramadorai, 2005), proximity and cultural similarities-induced home bias (Chan, Covrig and Ng, 2005), currency denomination (Maggiori, Neiman and Schreger, 2020), diversification and learning motives (Agarwal, Gu and Prasad, 2020), information and transaction technology (Portes and Rey, 2005), tax haven status (Coppola, Maggiori, Neiman and Schreger, 2021), and economic uncertainty (Alok, Javadekar, Kumar and Wermers, 2022).

[6] The literature on capital flows to emerging markets documents the role of interest rate differentials, U.S. monetary policy, and global risk aversion (Lee and Engel, 2024; Ahmed and Zlate, 2014; Ghosh, Qureshi, Kim and Zalduendo, 2014; Forbes and Warnock, 2012; Hutchison and Noy, 2006).

country in the sample, we generate time series measures that capture both the volume of China-related media coverage at the extensive margin and the sentiment of that coverage at the intensive margin. In addition, we construct topic-level attention and sentiment indices that measure the prominence of specific topics in investor-country media reporting about China and the sentiment associated with those topics.

2.1 Data Collection

We include media outlets, typically newspapers, in countries where institutional investors maintain significant holdings of Chinese assets. For English-speaking countries, our focus is on newspapers that are most widely circulated. In non-English-speaking countries, we prioritize the leading English-language newspapers. Our news article data are sourced from ProQuest TDM Studio, which also guided our final selection of media outlets based on the available sources within the platform (see Table A1 in Appendix A for a full list). In addition, we include the *Financial Times*, both as one of the U.K. media outlets and as the primary media source for European Economic and Monetary Union (EMU) countries excluding Ireland.[7]

To identify news articles related to China, we conducted searches using the keyword "China" for each media outlet. For the Taiwan-based *China Post*, and the Hong Kong-based *China Daily (Hong Kong ed.)*, both of which include "China" in their names, we used the keyword "mainland China" for our searches.[8]

If TDM Studio fails to collect content for a newspaper for more than five days within a single month, we exclude the data for that entire month from our analysis. Additionally, if there is a significant deviation in the volume of news articles for a given month relative

[7]We define the EMU following Maggiori, Neiman and Schreger (2020) and Coppola, Maggiori, Neiman and Schreger (2021). The EMU countries in their data include Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, the Netherlands, Portugal, Slovenia, and Spain. Each country enters the sample only after adopting the euro: Malta in 2008, Slovenia in 2007, and the remaining countries in 2002.

[8]As a robustness check, we also identify China-related news by assigning geographic location tags to each article using spaCy and selecting articles associated with China-related locations. A comparison based on a selected sample indicates that discrepancies between the two approaches are minimal.

to adjacent months, indicating a significantly lower count, the data for such a month are omitted. These adjustments help maintain the integrity and consistency of the dataset by eliminating periods of data scarcity or potential reporting anomalies. Detailed information on the sample period for each media outlet can be found in Table A1 in Appendix A. The final dataset includes 1,484,526 China-related news articles from 39 newspapers across 16 economies, covering the period from January 1, 2007 to May 31, 2022.

To mitigate potential language bias arising from the exclusive use of English-language newspapers, we incorporate seven additional newspapers in German, French, and Spanish, covering countries including Germany, France, Spain, Switzerland, and Austria. Table A2 reports the set of non-English newspapers included in this extended analysis.

2.2 Index Construction

Using natural language processing method, we construct four key narrative indices: an attention index, an aggregate sentiment index, a topic-level share index, and a topic-level sentiment index. At the extensive margin, the attention index, denoted by $num_{c,t}$, measures the volume of China-related news and quantifies the overall level of media attention to China. At the intensive margin, we compute an average sentiment index, $sen_{c,t}$, at the article level, which captures changes in the tone of reporting about China. We further construct attention and sentiment indices using non-English newspapers as a robustness check. Using large-language models, we further decompose China-related news into twelve thematic topics and construct corresponding indices at the topic level. The topic share index, $share^k_{c,t}$, measures the proportion of China-related news devoted to topic k relative to all China-related news in a given country and month, capturing how media attention is allocated across topics. The topic sentiment index, $sen^k_{c,t}$, measures the average sentiment associated with topic k, reflecting how positively or negatively that topic is portrayed. Taken together, these measures provide a comprehensive characterization of both the breadth and the intensity of media narratives about China across countries and

over time.

2.2.1 Aggregate-level Index

Attention Index. We follow a methodology similar to Baker, Bloom and Davis (2016) to construct the attention index, $num_{c,t}$, for investors domiciled in country c in month t, which serves as a measure of China related news volume at the extensive margin. For each media outlet m, we scale the number of China related articles published in month t by the total number of articles published by the same outlet in that month. This normalization controls for differences in overall news output across media outlets. We then compute the country level attention index by averaging these scaled measures across all media outlets within each country.

Sentiment Index. We construct a sentiment index, $sen_{c,t}$, for each country c in month t using the FinBERT methodology, which applies transformer-based sentiment classification models pre-trained or fine-tuned on financial text. For the English-language analysis, we use the `ProsusAI/finbert` model, an industry-standard BERT variant trained on financial news and earnings call transcripts. A detailed description of the model and implementation is provided in Appendix D. Examples of article-level classifications are reported in Table A3.

The model classifies text into three sentiment categories, positive, neutral, and negative, each accompanied by a confidence score. We apply the classifier to a truncated version of each article consisting of the headline followed by the first 480 characters of the body text. This window captures the portion of the article where narrative framing and evaluative language are typically concentrated, while avoiding dilution from descriptive detail or repetition later in the article.[9] This approach preserves the most sentiment-relevant

[9] BERT-based models, including FinBERT, operate under fixed input length constraints, and the average article in our dataset contains approximately 695 words. Beyond this mechanical consideration, truncation reflects a deliberate design choice. Restricting attention to the headline and opening text ensures consistent treatment across outlets and languages and aligns with journalistic conventions, in which the headline and

content while maintaining comparability across a large corpus of documents and multiple language settings.

After extracting article-level sentiment labels from the `ProsusAI/FinBERT` model, we convert qualitative classifications into a numeric time series by assigning a value of +1 to positive articles, 0 to neutral articles, and −1 to negative articles. Each article-level sentiment value is multiplied by the model-assigned confidence score, so that classifications made with greater certainty receive proportionally more weight. We then aggregate to the newspaper-month level by taking the mean across all articles published by a given newspaper within a month. To obtain a country-level series, we further average the monthly newspaper values across all newspapers available for that country. The resulting monthly sentiment index reflects the weighted balance of positive and negative language in national media coverage and is comparable across countries.

Table 1 reports summary statistics for the sentiment index constructed using the FinBERT method ($sen_{c,t}$). The correlation between this FinBERT-based index and the traditional bag-of-words–based sentiment index is 0.82, indicating a high degree of consistency across methods.[10]

Cross-Language Sentiment Index. In our baseline construction, we rely on English-language newspapers for the natural language processing analysis to ensure direct comparability across countries, as cross-language differences in vocabulary, morphology, and contextual usage could otherwise complicate the interpretation of sentiment indices. At the same time, the FinBERT-based methodology provides a natural framework for extending the analysis beyond English. Transformer-based sentiment models can be fine-tuned separately for different languages while preserving a common polarity structure, namely `positive`, `neutral`, and `negative`, which allows us to incorporate non-English text in a manner that remains systematically comparable across countries. This extension is

lead paragraph convey the most salient sentiment and narrative framing.

[10]We further construct a sentiment index using bag-of-words–based method and a risk index. The methodology is reported in Appendix B.

particularly relevant for European economies, where leading national newspapers are typically published in local languages and where key economic and political narratives may be more accurately reflected in domestic-language media.

Implementing this multilingual extension requires selecting appropriate FinBERT-style models for each language. We evaluate a range of candidates for German, French, and Spanish, focusing on domain alignment, calibration quality, and consistency with the FinBERT labeling framework. Appendix D documents the model-selection process and reports the final packages adopted for each language. Because these models share the same polarity structure as the English FinBERT, they enable a unified narrative-measurement framework across countries and languages.

We include seven additional newspapers in German, French, and Spanish, covering Germany, France, Spain, Switzerland, and Austria. We construct sentiment indices from these sources and incorporate them into our empirical analysis. Importantly, this multilingual extension is not merely a robustness exercise but also constitutes a methodological contribution of independent interest. By comparing sentiment derived from English-language reporting with sentiment extracted from dominant domestic-language news ecosystems, we directly assess whether cross-country variation in our baseline indices is driven by linguistic or editorial differences or instead reflects deeper structural differences in national media narratives.

2.2.2 Topic-level Index

We further classify all news articles in our dataset into twelve mutually exclusive subject categories using a large language model. These categories capture the primary dimensions along which global media report on China and provide a structured representation of the thematic content embedded in news coverage. The twelve topics are defined as follows:

1. **Trade & Supply Chain:** Exports, imports, tariffs, sanctions, supply-chain disruptions, logistics, commodity trade, and sourcing.

2. **Domestic Economy & Growth:** GDP, inflation, consumption, investment, unemployment, industrial production, macroeconomic performance, and stimulus policies.

3. **Financial Markets, Banking & FX:** Stocks, bonds, currencies, interest rates, central banking, monetary policy, capital flows, and banking developments.

4. **Real Estate, Debt & Financial Stability:** Property markets, mortgages, developers, credit risks, and broader financial fragility concerns.

5. **Technology, Innovation & Industry Policy:** Technology firms, semiconductors, artificial intelligence, telecommunications, industrial upgrading, automation, and state-led industrial policy initiatives.

6. **Governance, Regulation & Institutions:** Government policy actions, regulatory interventions, administrative measures, leadership decisions, and institutional reforms.

7. **Environment, Climate & Energy:** Pollution, climate change, renewable energy, fossil fuels, mining, and natural resource issues.

8. **Diplomacy & Geopolitics:** Foreign relations, great-power competition, military relations, alliances, interstate conflicts, and China's positioning in global politics.

9. **Society, Labor & Demographics:** Population trends, migration, education, workforce issues, inequality, and civil-society developments.

10. **Security, Surveillance & Human Rights:** National security, policing, censorship, surveillance technologies, detentions, and human-rights disputes.

11. **Health, Pandemics & Public Safety:** Disease outbreaks, vaccines, COVID-19, medical systems, and public-safety concerns.

12. **Firms, Industries & Corporate Strategy:** Corporate decisions, earnings, mergers and acquisitions, product launches, supply–demand conditions, and business strategies.

Instead of relying on traditional unsupervised approaches such as Latent Dirichlet Allocation (LDA), which often produce unstable clusters and require subjective interpretation, we classify articles using a pre-defined taxonomy that aligns with key economic, political, and social domains. The classification is performed using the `gpt-4.1-mini` model, which interprets headlines with high semantic accuracy and applies the taxonomy consistently. For each article, only the headline is provided to the model together with a concise description of the twelve categories, and the model returns a single topic label. This procedure yields a transparent and replicable mapping from raw text to economically

meaningful subjects.[11] The exact prompt and the full category definitions are reported in Appendix F.

For the *Financial Times* (FT) archive, we follow the same taxonomy and classification procedure, but due to data usage restrictions,[12] classification is implemented via a local deployment of the `mistral-7b-instruct.Q4_0` model using the `llama.cpp` inference framework. While smaller than commercial cloud models, this architecture offers fast inference, robust semantic understanding of headlines, and full compatibility with constrained environments, making it a practical and secure solution for large-scale, title-based classification. To mitigate potential misclassification arising from the limitations of a local model, we use both the article title and the first two sentences of each *Financial Times* article when assigning subject classifications.

Using an LLM provides several methodological advantages relative to LDA. Because the taxonomy is fixed and conceptually grounded, topic assignments remain stable across newspapers, time periods, and subsamples, whereas LDA topics change as corpus composition evolves. The LLM also incorporates contextual and domain-specific knowledge that improves semantic differentiation, allowing it to distinguish, for example, macroeconomic developments from financial market interventions or diplomatic communication from national security concerns, distinctions that LDA often merges into broad clusters. The procedure also scales efficiently: classifying approximately 180,000 Wall Street Journal articles required about four hours and a modest monetary cost, making the approach feasible for our multi-million–article dataset. Since all articles are classified under the same taxonomy, the resulting subject indices are directly comparable across countries and over time, which is essential in an international setting. Figure 5 presents bigram word clouds for each topic, demonstrating that the classification method effectively captures the

[11] To reduce computational time and cost, we classify articles using only their titles rather than full text. Random subsample tests confirm that title-based and full-text classifications produce near identical results.

[12] The FT license agreement limits how content may be used in conjunction with AI systems. Specifically, classification must be conducted on locally hosted models entirely controlled by the user, and results cannot be used to train or adapt any general-purpose large language model. The full content cannot be transferred to third-party cloud models or services.

core thematic content within each category.

Formally, for each topic k, country c, and month t, we construct a topic share index, $share^k_{c,t}$, and a topic sentiment index, $sen^k_{c,t}$, for each country c at time t.

Topic Share Index. The topic share index, $share^k_{c,t}$, measures the proportion of media attention devoted to topic k in country c and month t relative to all China related news. For each newspaper, we compute the share of articles classified as topic k in month t relative to the total number of China related articles published by that newspaper in the same month. This index captures how reporting capacity is allocated to topic k out of all China related coverage. We then average these newspaper level shares across all newspapers in country c to obtain the monthly index.

Topic Sentiment Index. The topic sentiment index $sen^k_{c,t}$ captures the intensive margin of narrative tone for topic k in country c in month t. For each newspaper, we compute the average FinBERT sentiment score of articles classified as topic k in that month, which reflects the within-subject tone conditional on coverage. We then take the average of these values across all newspapers in country c to obtain the monthly index.

3 Stylized Facts

3.1 Visualizing the Indices

Figure 1 shows that both the attention and sentiment indices track major global and China-related episodes closely, confirming that our narrative measures capture systematic shifts in international reporting. Attention to China rises sharply around well-known events, including the Global Financial Crisis, the 2015–2016 China stock market crash, the escalation of U.S.-China trade tensions, and the onset of the COVID-19 pandemic, while sentiment turns markedly more negative during these periods. Notably, the cross-country interquartile range narrows during these major episodes, indicating that media outlets

across different economies tend to converge in both the volume and tone of their China coverage when global uncertainty spikes. Outside of crisis periods, however, the dispersion widens again, reflecting persistent cross-country heterogeneity in how China is framed in the news.

Figure 3 presents the evolution of topic-level media narratives about China between 2007 and 2022. Both the topic share and sentiment indices display substantial time variation, with noticeable increases in coverage and shifts in tone during major global or China-related events. For example, attention to Health, Pandemics & Public Safety (Topic 11) rises sharply around 2020, consistent with the outbreak of COVID-19, while coverage of Financial Markets, Banking & FX (Topic 3) spikes during the 2015–2016 stock market turmoil. Sentiment patterns similarly reflect key episodes: sentiment within economic and financial topics deteriorates during the Global Financial Crisis and again during the 2018–2020 trade tensions, whereas technology-related sentiment (Topic 5) remains relatively stable over time. Across topics, the interquartile range shows considerable cross-country heterogeneity, though dispersion narrows temporarily during large global shocks when reporting becomes more synchronized.

3.2 Cross-Country Narrative Dispersion

Topic Share Dispersion. Across countries, the topical composition of China-related reporting exhibits clear and systematic patterns. Panel (a) of Figure 4 shows that three topics, Diplomacy & Geopolitics (Topic 8), Society, Labor & Demographics (Topic 9), and Firms, Industries & Corporate Strategy (Topic 12), receive consistently higher coverage shares than other categories across most economies. These topics account for a substantial portion of China-related reporting, indicating that global media outlets tend to frame China primarily through geopolitical relations, domestic social developments, and corporate or industrial dynamics.

At the same time, the heatmap reveals notable cross-country clustering in narrative

emphasis. Economies geographically close to China, such as Korea, India, Thailand, and Taiwan, devote relatively more coverage to Diplomacy & Geopolitics (Topic 8), reflecting the salience of regional security and bilateral relations. Advanced western economies, including the United States, United Kingdom, and Canada, allocate disproportionately large shares to Diplomacy & Geopolitics (Topic 8) and Society, Labor & Demographics (Topic 9), consistent with a more political and socio-institutional framing of China. By contrast, Asian economies with deeper commercial exposure to China, such as Hong Kong, Malaysia, Singapore, and Korea, place greater weight on Firms, Industries, & Corporate Strategy (Topic 12), and Financial Markets, Banking & FX (Topic 3), indicating a more economic and business-oriented lens on China's global role.

Topic Sentiment Dispersion. Panel (b) of Figure 4 highlights pronounced cross-topic and cross-country heterogeneity in sentiment toward China. First, there are strong systematic differences across topics: Security, Surveillance & Human Rights (Topic 10) is uniformly negative across all countries, and especially so in the United States, United Kingdom, Canada, Australia, and New Zealand, making it the single most negative narrative category. Governance (Topic 6) and Health & Public Safety (Topic 11) also exhibit consistently negative sentiment across the sample. By contrast, Firms, Industries & Corporate Strategy (Topic 12) and Technology, Innovation & Industry Policy (Topic 5) are among the few categories with generally positive sentiment, with particularly favorable views in Malaysia, the Philippines, and Thailand, although Topic 12 shows notable exceptions for the United States and United Kingdom, where sentiment is markedly more negative.

Second, several topics exhibit substantial cross-country disagreement in tone. Sentiment toward Domestic Economy & Growth (Topic 2), Financial Markets, Banking & FX (Topic 3), and Real Estate, Debt & Financial Stability (Topic 4) varies widely across countries. Southeast Asian economies, including Malaysia, the Philippines, and Thailand, tend to report these topics with a more positive tone, whereas advanced Western economies

report them much more negatively, indicating divergent assessments of China's economic conditions.

Third, when comparing sentiment across all topics, a clear geographic divide emerges: Asian economies such as Malaysia, the Philippines, and Thailand consistently display more positive sentiment toward China, while the United States, United Kingdom, Canada, and Australia are systematically more negative. This persistent cross-country asymmetry highlights distinct national narrative environments, with implications for how investors in different countries interpret China-related information.

4 Understanding Cross-Country Narrative Dispersion

We further examine the sources of cross-country dispersion in media narratives about China. While aggregate sentiment toward China varies substantially across investor countries, such variation can arise for conceptually distinct reasons. Countries may differ in the extent to which their media devote attention to different aspects of China, what we refer to as *topic attention* (or topic coverage), or they may differ in how they evaluate and frame the same topics, as reflected in within-topic sentiment. Distinguishing between these channels is important for understanding whether cross-country narrative disagreement primarily reflects differences in information selection or differences in interpretation. To this end, we develop a variance decomposition of aggregate sentiment that isolates the contribution of topic attention and within-topic sentiment to cross-country dispersion.

4.1 Topic Attention versus Within-Topic Sentiment

We conduct a counterfactual variance decomposition to separate variation arising from differences in topic attention, i.e., the allocation of media attention across certain topic, from variation arising from differences in sentiment conditional on topic. We do so by constructing counterfactual sentiment indices that hold one component fixed at its cross-country mean while allowing the other to vary across countries. Comparing the cross-

country variance of these counterfactual indices to the variance of observed aggregate sentiment allows us to quantify the relative contribution of topic attention, within-topic sentiment, and their interaction to overall narrative disagreement.

Let $S_{c,t}$ denote the aggregate sentiment toward China in country c at time t, constructed as a weighted average of topic-level sentiment, where weights reflect topic attention. Formally,

$$S_{c,t} = \sum_{k=1}^{K} w_{c,t}^{k} s_{c,t}^{k},$$

where $w_{c,t}^{k}$ denotes the share of media coverage devoted to topic k in country c at time t, corresponding to the topic share index $share_{c,t}^{k}$, and $s_{c,t}^{k}$ denotes sentiment toward topic k in country c at time t, corresponding to the topic sentiment index $sen_{c,t}^{k}$.

To isolate cross-country variation in sentiment $S_{c,t}$ driven purely by differential topic attention across investor countries' media, we construct a counterfactual sentiment index that holds within-topic sentiment fixed at its cross-country mean. Formally,

$$S_{c,t}^{\text{att}} = \sum_{k=1}^{K} w_{c,t}^{k} \bar{s}_{t}^{k}, \qquad \bar{s}_{t}^{k} \equiv \frac{1}{C} \sum_{c=1}^{C} s_{c,t}^{k}.$$

where $w_{c,t}^{k}$ denotes the share of media attention devoted to topic k in country c at time t, and $\bar{s}_{t}^{k}$ is the cross-country average sentiment for topic k at time t. By construction, this index varies across countries solely because of differences in topic attention, abstracting from cross-country differences in sentiment conditional on topic.

Conversely, holding topic attention fixed at its cross-country mean, we define the within-topic-sentiment-only counterfactual index as

$$S_{c,t}^{\text{tone}} = \sum_{k=1}^{K} \bar{w}_{t}^{k} s_{c,t}^{k}, \qquad \bar{w}_{t}^{k} \equiv \frac{1}{C} \sum_{c=1}^{C} w_{c,t}^{k}.$$

where $\bar{w}_{t}^{k}$ is the cross-country average topic attention weight for topic k at time t, and

$s^k_{c,t}$ denotes country-specific sentiment toward topic k. This index varies across countries solely because of differences in sentiment conditional on topic, holding the allocation of attention across topics fixed.

Next, we study the relative contribution of the two counterfactual indices to the variance of the country-level sentiment index. Formally, let $\mathrm{Var}_t(\cdot)$ denote the cross-country variance at a given time t. We compute:

$$V_t = \mathrm{Var}_t(S_{c,t}); \qquad V_t^{\text{att}} = \mathrm{Var}_t\left(S_{c,t}^{\text{att}}\right); \qquad V_t^{\text{tone}} = \mathrm{Var}_t\left(S_{c,t}^{\text{tone}}\right).$$

For each time period t, V_t measures the total cross-country dispersion in aggregate sentiment toward China, while V_t^{att} and V_t^{tone} measure the portions of that dispersion generated by cross-country differences in topic attention and within-topic sentiment, respectively, as captured by the corresponding counterfactual sentiment indices.

Using these quantities, we define the variance shares attributable to topic attention and within-topic sentiment as

$$\text{Share}_t^{\text{att}} = \frac{V_t^{\text{att}}}{V_t}, \qquad \text{Share}_t^{\text{tone}} = \frac{V_t^{\text{tone}}}{V_t}.$$

These shares measure the fraction of cross-country variation in aggregate sentiment at time t that can be attributed to differences in topic attention and differences in sentiment conditional on topic, respectively.

Because topic attention and within-topic sentiment may not be independent across countries, the variation in aggregate sentiment will also depend on the interaction between the two components. We therefore define a residual interaction component as

$$\text{Residual}_t = 1 - \text{Share}_t^{\text{att}} - \text{Share}_t^{\text{tone}}, \tag{1}$$

which captures cross-country variation arising from the interaction between topic attention

and within-topic sentiment.

Figure 6 summarizes the results of the variance decomposition. We find that the overwhelming majority of cross-country dispersion in aggregate sentiment toward China is driven by differences in within-topic sentiment rather than by differential topic coverage across countries. In contrast, variation in topic attention contributes only a small fraction of overall narrative disagreement. This finding indicates that cross-country differences in media narratives about China do not primarily reflect what countries choose to write about, but rather how they evaluate and frame the same underlying topics. Put differently, investor-country media largely focus on similar sets of China-related issues, yet attach systematically different sentiment to those issues. As a result, disagreement in aggregate narratives arises mainly from differences in tone conditional on topic, rather than from differences in the allocation of attention across topics.

4.2 Heterogeneous Belief Levels versus Heterogeneous Updating

While the preceding analysis establishes that cross-country narrative dispersion is driven primarily by within-topic differences in sentiment, it does not yet explain why narratives diverge within the same topic. Conceptually, within-topic sentiment can differ across countries through two channels. First, there is belief-level heterogeneity, which captures differences in countries' baseline framing of a given topic. This channel itself has two components. On the one hand, countries may enter the sample with systematically different baseline stances toward China within a topic, what we interpret as heterogeneous priors. These differences generate persistent dispersion even in the absence of new information. On the other hand, even when exposed to the same topic-level information, countries may systematically translate that information into sentiment differently, reflecting heterogeneous interpretation of common signals. Second, there is updating heterogeneity, whereby countries differ in how strongly they adjust sentiment in response to new information relative to their existing beliefs. That is, countries may place different weights on prior stances

versus common topic-level signals, generating cross-country differences in updating intensity. In this section, we decompose within-topic sentiment into persistent country-specific components and a country-specific response to common topic-level signals.

4.2.1 Bayesian learning model

We develop a Bayesian learning framework that decomposes cross-country variation in media narratives about China into differences in persistent country-specific priors and differences in how countries translate common information into sentiment following Lahiri and Sheng (2008). For each topic k, investor country c, and month t, let $S^k_{c,t}$ denote the media sentiment index capturing country c's narrative stance toward China on topic k, corresponding to the topic sentiment index $sen^k_{c,t}$ defined in the previous section. We interpret $S^k_{c,t}$ as a proxy of the posterior mean belief about a latent topic-k "state" of China that is updated as new public information arrives.

Latent state, priors, and public information. Let θ^k_t denote the latent topic-k state of China at month t. Country c enters month t with a (Gaussian) prior belief about θ^k_t:

$$\theta^k_t \,\big|\, \mathcal{I}_{c,t-1} \sim \mathcal{N}\Big(S^k_{c,t-1},\, (a^k_{c,t-1})^{-1}\Big), \tag{2}$$

where $S^k_{c,t-1}$ is the prior mean and $a^k_{c,t-1}$ is the prior precision.[13]

At month t, all countries observe a common public signal L^k_t about θ^k_t but may interpret it differently. Specifically, country c's newspapers' (subjective) estimate of θ^k_t based only on the new public signal as

$$Y^k_{c,t} = L^k_t - \varepsilon^k_{c,t}, \qquad \varepsilon^k_{c,t} \sim \mathcal{N}\Big(\mu^k_{c,t},\, (b^k_{c,t})^{-1}\Big), \tag{3}$$

[13] We allow the precision of initial prior beliefs about China to vary across countries, reflecting heterogeneity in information uncertainty. Such differences may arise from country-specific characteristics, including geographic distance, time zone differences, cultural proximity, media infrastructure, etc. Similar assumptions of heterogeneous prior precision are standard in the forecasting literature (e.g., Lahiri and Liu, 2006; Lahiri and Sheng, 2008).

where $\mu_{c,t}^k$ is an *interpretation bias*, which is the mean wedge in mapping the common signal into beliefs, and $b_{c,t}^k$ is the signal precision, which captures "latitude" in interpretation.

Posterior mean updating rule Under Gaussian priors and signals, Bayes' rule implies the posterior mean belief is a convex combination of the prior mean and the (bias-adjusted) signal:

$$S_{c,t}^k = \lambda_{c,t}^k S_{c,t-1}^k + (1 - \lambda_{c,t}^k)\Big(L_t^k - \mu_{c,t}^k\Big), \tag{4}$$

where the *weight on the prior* is

$$\lambda_{c,t}^k \equiv \frac{a_{c,t-1}^k}{a_{c,t-1}^k + b_{c,t}^k} \in (0,1). \tag{5}$$

The object $(1 - \lambda_{c,t}^k)$ is the *weight on new public information* (updating intensity).

4.2.2 Descriptive Analysis: Belief-Level Heterogeneity and Updating Heterogeneity.

For descriptive (static) cross-country plots, we estimate time-invariant country–topic parameters (λ_c^k, μ_c^k) implied by the Bayesian updating rule (4). Because the common public signal L_t^k is not directly observed, we proxy it with the topic-k global information index constructed in the paper using principal components.

Leave-one-out proxy for the common public signal. To mitigate mechanical correlation between the signal proxy and a given country's sentiment series, we construct a *leave-one-out* topic-k information factor for each country c:

$$\widehat{L}_{-c,t}^k \;\equiv\; \mathrm{PC1}_t\Big(\{S_{c',t}^k\}_{c' \neq c}\Big), \tag{6}$$

i.e., the first principal component extracted from the panel of topic-k sentiment indices excluding country c. For each (c,k), we normalize $\widehat{L}_{-c,t}^k$ to have mean zero and unit variance over t. Because principal components are identified only up to sign, we orient $\widehat{L}_{-c,t}^k$ so that

it has positive correlation with the cross-country mean $\bar{S}_t^k$.

Country-specific interpretation wedges and updating weights. Imposing time-invariant country–topic parameters $(\lambda_{c,t}^k, \mu_{c,t}^k) = (\lambda_c^k, \mu_c^k)$ and replacing L_t^k with $\widehat{L}_{-c,t}^k$ in (4), subtracting $\widehat{L}_{-c,t}^k$ from both sides yields:

$$\begin{aligned} S_{c,t}^k - \widehat{L}_{-c,t}^k &= \lambda_c^k \left(S_{c,t-1}^k - \widehat{L}_{-c,t}^k \right) - (1-\lambda_c^k)\mu_c^k + u_{c,t}^k \\ &\equiv \delta_c^k + \lambda_c^k \left(S_{c,t-1}^k - \widehat{L}_{-c,t}^k \right) + u_{c,t}^k, \end{aligned} \tag{7}$$

where $\delta_c^k \equiv -(1-\lambda_c^k)\mu_c^k$ is a country–topic intercept and $u_{c,t}^k$ collects measurement error in $\widehat{L}_{-c,t}^k$ and any deviations from time-invariant (λ_c^k, μ_c^k).

Estimation. We estimate (7) separately by topic k on the pooled (c,t) panel, allowing for country-specific λ_c^k and δ_c^k via interactions with country indicators. Standard errors are clustered at the country level, and we impose $\widehat{\lambda}_c^k \in [0, 0.95]$ for numerical stability. We then recover the implied interpretation wedge and updating intensity:

$$\widehat{\mu}_c^k \equiv -\frac{\widehat{\delta}_c^k}{1-\widehat{\lambda}_c^k}, \qquad \widehat{\kappa}_c^k \equiv 1 - \widehat{\lambda}_c^k. \tag{8}$$

Visualization: belief Level heterogeneity and updating heterogeneity. Figure 7 visualizes the time-invariant country–topic parameters implied by the Bayesian updating rule in (4). Panel (a) reports the estimated interpretation wedge $-\widehat{\mu}_c^k$ (higher values indicate a *more favorable* mapping from the common topic signal into sentiment), while Panel (b) reports the updating intensity $\widehat{\kappa}_c^k \equiv 1 - \widehat{\lambda}_c^k$, i.e., the average weight placed on new public information. Two patterns stand out. First, cross-country differences are far more pronounced in *belief-level* objects than in *updating*: $-\widehat{\mu}_c^k$ exhibits large and easily visible dispersion across countries within most topics, whereas $\widehat{\kappa}_c^k$ is tightly concentrated near zero with only a handful of darker "hot spots." This contrast implies that international differences in

China-related media sentiment are driven primarily by how countries *interpret* common information (systematic wedges) rather than by large differences in the average speed at which they incorporate new information.

Second, the sign and magnitude of interpretation wedges vary systematically across topics and country groups. Topics related to Security, Surveillance & Human Rights (Topic 10) and Health, Pandemics & Public Safety (Topic 11) display broadly negative interpretation wedges across many countries, while Technology, Innovation & Industry Policy (Topic 5) and Firms, Industries & Corporate Strategy (Topic 12) are comparatively more positive. Cross-country clustering is also evident: Malaysia and the Philippines tend to exhibit more favorable wedges across *most* topics (albeit less so in Topics 10–11), while the U.S., the U.K., Australia, and Canada are generally more negative, consistent with persistent differences in narrative framing. In the next section, we further analyze the origins of those national belief levels.

By contrast, updating intensities vary less across countries, but the largest $\widehat{\kappa}_c^k$ values appear in information-rich domains, notably Domestic Economy & Growth (Topic 2), Financial Markets (Topic 3), and Real Estate / Financial Stability (Topic 4), and (for some countries) Health (Topic 11), with especially high updating for Taiwan and comparatively higher updating in a subset of Asia-Pacific and emerging economies (e.g., Malaysia, New Zealand, and South Africa), while the advanced economies such as the U.S., the U.K., Canada, and Australia, exhibit consistently low $\widehat{\kappa}_c^k$ (high $\widehat{\lambda}_c^k$), indicating greater inertia in how common news translates into sentiment.

4.2.3 Variance Decomposition: Belief-Level and Updating Heterogeneity

Variance Decomposition. The Bayesian updating model yields the following tractable recursion for disagreement:

$$\sigma^2_{S|t,k} = \sigma^2_{S|t-1,k}\Big(\sigma^2_{\lambda|t,k} + (\lambda_t^k)^2\Big) + \sigma^2_{\mu|t,k}\Big(\sigma^2_{\lambda|t,k} + (1-\lambda_t^k)^2\Big) + \sigma^2_{\lambda|t,k}\left[\frac{\Delta \bar{S}_t^k}{1-\lambda_t^k}\right]^2. \tag{9}$$

where

$$\bar{S}_t^k \equiv \mathbb{E}_c[S_{c,t}^k], \quad \Delta\bar{S}_t^k = \bar{S}_t^k - \bar{S}_{t-1}^k, \quad \sigma^2_{S|t,k} \equiv \mathrm{Var}_c(S_{c,t}^k), \quad \lambda_t^k \equiv \mathbb{E}_c[\lambda_{c,t}^k], \quad \sigma^2_{\lambda|t,k} \equiv \mathrm{Var}_c(\lambda_{c,t}^k).$$

Proof. Please see the appendix G for the proofs. ■

Equation (9) decomposes narrative disagreement into two components: First, Belief-level Heterogeneity: This is driven by (i) heterogeneity in inherited priors **P** (the term proportional to $\sigma^2_{S|t-1,k}$), which propagates initial cross-country dispersion forward when narratives are inert, and (ii) interpretation of public information **I** ($\sigma^2_{\mu|t,k}$), which captures persistent cross-country dispersion in how common topic-level information is translated into sentiment;[14] Second, weights on priors **W** ($\sigma^2_{\lambda|t,k}$), which captures heterogeneity in updating intensity and is amplified during periods of large common mean revisions. Because equation (9) replaces the unobserved signal term with the *observable* mean revision $\Delta\bar{S}_t^k$, the variance decomposition does not require directly observing L_t^k.

Empirical implementation. We estimate the Bayesian-learning parameters topic by topic using the panel $\{S_{c,t}^k\}$. First, we recover the average updating weight λ^k from a forecast-revision regression augmented with common correlated effects controls (Pesaran, 2006). We then estimate the cross-country dispersion parameters $(\sigma^2_{\lambda|k}, \sigma^2_{\mu|k})$ from the regression representation implied by the variance recursion. To quantify the relative importance of the three channels $\{P, I, W\}$, belief-level heterogeneity (priors and interpretation) and updating heterogeneity, we generate counterfactual paths of cross-country disagreement in which selected channels are shut down while holding the others active. We then assess how closely each counterfactual path tracks the observed time series of disagreement by regressing observed disagreement on counterfactual disagreement and computing the

[14] In the previous descriptive analysis, the pooled panel specification identifies persistent country–topic interpretation wedges but does not separately identify inherited initial priors. As a result, the estimated belief-level heterogeneity primarily reflects differences in interpretation rather than explicitly modeled initial prior dispersion.

associated R^2. The marginal contribution of a given channel is measured by the change in explanatory power when that channel is removed. For example, the difference in R^2 between the full model and a specification in which channel I is shut down provides a measure of the contribution of heterogeneous interpretation to cross-country dispersion. We report both a ΔR^2 attribution and an order-invariant Shapley decomposition, which averages marginal contributions across all possible inclusion orders. All derivations, identifying assumptions, and estimation details are provided in Appendix G.

Results. Figure 8 decomposes cross-country dispersion in topic sentiment into (i) *belief-level disagreement*, inherited baseline stance differences (P) and heterogeneous interpretation of common public information (I); and (ii) *updating disagreement* (W), which captures cross-country heterogeneity in how strongly prior beliefs are reweighted relative to new information. Two complementary allocations (ΔR^2 and the order-invariant Shapley method) deliver the same qualitative message: belief-level components carry most of the variance across topics (on average about $P+I \simeq 70\%$ under ΔR^2 and $\simeq 61\%$ under Shapley), consistent with models in which heterogeneous priors and differential interpretation can generate persistent disagreement even under shared news (Cyert and DeGroot, 1974; Townsend, 1983; Haltiwanger and Waldman, 1985; Bray and Savin, 1986; Varian, 1989; Harris and Raviv, 1993; Kandel and Pearson, 1995; Kandel and Zilberfarb, 1999). At the same time, the updating channel becomes comparatively more important precisely where high-frequency, quantifiable signals are easiest to obtain, notably Topic 2 (Domestic Economy & Growth) and Topic 3 (Financial Markets, Banking & FX), for which W rises to 50% and 42% in the ΔR^2 decomposition (and 65% and 55% under Shapley), indicating that disagreement in these domains is disproportionately driven by heterogeneous reweighting of priors in response to incoming information. By contrast, for topics where information is less timely, less transparent, or intrinsically harder to map into a single "market" signal, such as Topic 1 (Trade & Supply Chain), Topic 4 (Real Estate, Debt &

Financial Stability), and Topic 8 (Diplomacy & Geopolitics), belief-level forces dominate (e.g., $P{+}I \simeq 79\%/87\%/78\%$ in LS and $\simeq 79\%/85\%/76\%$ in Shapley), suggesting that entrenched narratives and interpretive lenses ("stereotypes" and framing) have greater scope to persist when signal quality is uncertain (Acemoglu et al., 2006). These trends are also in the spirit of sticky-information and rational-inattention mechanisms (Mankiw and Reis, 2002; Sims, 2003). A further notable pattern is that I is often comparable to (and sometimes larger than) P, implying that persistent cross-country disagreement reflects not only different starting points but also systematic cross-country differences in how the same public information is read and incorporated into sentiment.

5 The Origins of National Belief Levels

In this section, we further investigate where cross-country differences in *time-invariant narrative belief level* about China come from. Let $\widehat{\mu}_c^k$ denote the estimated belief level for country c and topic k obtained in Section 4.2.2. We relate these belief levels to slow-moving (or time-invariant) country characteristics that proxy for economic interdependence, geopolitical alignment, proximity, institutional environment, and market depth and financial similarity.

Our specification pools all topics and includes topic fixed effects:

$$-\widehat{\mu}_c^k = \beta' X_c + \delta_k + \varepsilon_{c,k}, \tag{10}$$

where X_c is a vector of country characteristics and δ_k are topic fixed effects. Standard errors are two-way clustered at the country and topic levels.

We examine country-level determinants of cross-country belief dispersion following the literature on international heterogeneity (Head and Mayer, 2014; Bailey et al., 2017; Gurevich, 2018; Agarwal et al., 2020; Horn et al., 2021; Conte et al., 2022; Hassan et al.,

2024; Jiang, 2025).[15] For the variables, we use the average information from year 2007 to 2022 whenever available to be consistent with the time we used to estimate the belief level, otherwise we use the maximum time length available. Detailed definition of each of the variable can be found in Table A6 in Appendix. Specifically, we use the following variables:

1. **Economic interdependence with China.** We use bilateral trade intensity measures: (i) trade with China relative to GDP ($TradeCN_c$), and (ii) the share of the country in China's total trade ($TradeShare_c$), an "information endowment" proxy;

2. **Geopolitical alignment.** We proxy for geopolitical alignment using (i) UN voting alignment with China ($AgreeCN_c$) and (ii) Belt and Road Initiative participation (BRI_c);

3. **Proximity.** We use standard gravity-style measures: (i) Contiguity measure using common border with China ($ContCN_c$); (ii) Distance to China ($DistCN_c$); (iii) Common language ($ComLangCN_c$).

4. **Institutional quality.** We consider institutional quality measures: (i) rule of law ($RuleLaw_c$), (ii) government effectiveness ($GovEff_c$), and (iii) regulatory quality ($RegQuality_c$);

5. **Market depth and financial similarity.** We consider measures of market depth and financial similarity: (i) market capitalization (% of GDP) ($ListedFirms_c$), (ii) equity turnover ($Turnover_c$), and (iii) return-based similarity to China's stock market ($RetCorrCN_c$).

[15] A related literature studies media slant and ideological bias at the newspaper level (e.g., Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006; Besley and Prat, 2006; Gentzkow and Shapiro, 2010; Gentzkow et al., 2014; Qin et al., 2018). In our setting, because we aggregate across multiple major newspapers within each country, we do not expect idiosyncratic newspaper-level slant to drive cross-country narrative differences. Instead, we focus on systematic country-level determinants of narrative heterogeneity.

Results. Table 2 reports the cross-country determinants of narrative belief levels about China. Economic interdependence with China is positively associated with more favorable beliefs: both trade with China relative to GDP ($TradeCN_c$) and the share of bilateral trade in China's total trade ($TradeShare_c$) are positive and statistically significant in the parsimonious specification (column 8). Geopolitical alignment yields mixed effects: UN voting alignment with China is positively correlated with beliefs in the baseline alignment specification, while participation in the Belt and Road Initiative (BRI_c) is negatively associated with beliefs once other structural characteristics are controlled for. Geographic proximity also matters: countries sharing a border with China ($ContCN_c$) exhibit systematically different belief levels, whereas distance becomes insignificant once other covariates are included. Institutional quality plays a robust role, with higher rule of law associated with more negative belief levels in the full specification. Finally, financial integration appears economically meaningful: the return-based similarity measure ($RetCorrCN_c$) is strongly positive and highly significant, suggesting that countries whose equity markets co-move more closely with China tend to display more favorable narrative beliefs. By contrast, greater market depth, measured by the ratio of listed firms to GDP ($ListedFirms_c$) and equity turnover ($Turnover_c$), is generally associated with more negative beliefs. Overall, the results indicate that cross-country differences in belief levels reflect a combination of economic integration, geopolitical positioning, institutional structure, and financial linkage to China.

6 Local Media Narratives and Cross-Border Flows to China

Having established that there is substantial cross-country variation in media narratives, we now turn our attention to the impact of local media narratives in investors' domicile countries on cross-border investments in China. We formally investigate the extent to which media narratives shape portfolio flows to China, utilizing quarterly global institutional investor portfolio holdings data from Morningstar. The majority of foreign

holdings of Chinese assets are accounted for by open-ended funds, which also tend to be more active investors relative to ETFs and money market funds in response to market and other developments. Hence, our empirical analysis focuses on open-ended funds.

We estimate the impact of local media narratives on cross-border flows of institutional investors using the following regression specification:

$$Flow_{ic,t} = \beta_0 + \beta_1 Index_{c,t-1} + \beta_2 num_{c,t-1} + X'_{c,t-1}\Phi + \alpha_i + \alpha_t + \varepsilon_{ic,t} \tag{11}$$

where i indexes funds, c indexes fund domicile country, and t indexes calendar quarter. $Flow_{ic,t}$ measures fund i's capital flow to Chinese assets in quarter t. Specifically, we define flow as follows:

$$Flow_{ic,t} = \frac{\sum_{a \in C_{ic,t}} P^a_{t-1}(N^a_{ic,t} - N^a_{ic,t-1})}{\sum_{a \in C_{ic,t-1}} P^a_{t-1} N^a_{ic,t-1}}$$

where $C_{ic,t}$ is the collection of all unique Chinese assets held by fund i at quarter t domiciled in the country c. P^a_{t-1} denotes the dollar value of the asset a at the end of quarter $t-1$. $N^a_{ic,t}$ is the number of units of asset a held by fund i domiciled in c at quarter t. Therefore, the numerator represents the total value of changes in Chinese asset holdings, calculated using the unit price of the asset from the previous quarter, which excludes the impact of market price fluctuations.[16]

$Index_{c,t-1}$ denotes the media narrative based sentiment index for investors in domicile country c at time $t-1$, and serves as our main variable of interest. We also include $num_{c,t-1}$ to control for the effect of narrative coverage at the extensive margin on investment flows.[17] $X_{c,t-1}$ is a vector of macroeconomic controls, lagged one quarter, that may affect fund flows to China. These include the year-over-year GDP growth differential between China and the investor's domicile country ($Growthdif f_{c,t-1}$); the bilateral exchange rate

[16]Since it is hard to interpret valuation changes for derivative assets, we exclude all derivative holdings from our analysis. This ensures that our data reflects more accurately the actual market positions without the distortions that derivatives might introduce due to their complex accounting treatments. Additionally, to reduce the impact of extreme outliers, $Flow_{ic,t}$ has been winsorized at the 0.5% level in each tail.

[17]Since the indices are constructed at a monthly frequency, we use the arithmetic average over the three months of the quarter ending in $t-1$ in equation 11.

between the renminbi and the domicile country's currency ($EX_{c,t-1}$), where an increase reflects a depreciation of the renminbi; the interest rate differential between China and the domicile country ($Intdiff_{c,t-1}$); and relative equity market performance, measured by the excess return of the Chinese stock market over the domicile country's market ($Retdiff_{c,t-1}$); and the volatility of the Chinese stock market over the volatility of the domicile country's market ($Volratio_{t-1}$).[18] To facilitate cross-country comparisons, we standardize the sentiment and risk indices for each country. The specification includes fund fixed effects (α_i) and quarter-year fixed effects (α_t). Standard errors are clustered at the fund level to adjust for potential serial correlation in the error term $\varepsilon_{ic,t}$.

Table 3 reports baseline estimates from alternative specifications of equation (11) using different combinations of control variables. Columns (1) through (4) show that more positive domestic media sentiment toward China is associated with higher fund flows into Chinese assets. As shown in column (4), after controlling for macroeconomic and financial fundamentals, as well as fund and quarter fixed effects, a one standard deviation increase in the domestic media sentiment index is associated with a 0.96% increase in quarterly investment flows to China, which corresponds to an annualized increase of 3.82%.

The inclusion of quarter fixed effects ensures that these results are not driven by common global shocks, highlighting that cross-country differences in media sentiment play an important role in explaining heterogeneity in capital flows. Consistent with the earlier cross-country variance decomposition, which shows that most narrative variation arises from sentiment rather than attention, we find that the volume of China-related news coverage does not have a statistically significant relationship with investment flows once sentiment is controlled for. This result indicates that investment decisions respond primarily to the tone of media narratives at the intensive margin, rather than to changes in the overall quantity of China-related reporting.

[18] Definitions and data sources for all control variables are provided in Table A4 in Appendix E.

6.1 Robustness Analysis: Bag-of-word Method

We further construct bag-of-word sentiment indices ($sen_{c,t}^{bag}$) using the method described in Appendix B, in which we count the number of positive and negative words in each news article based on the word lists developed by Loughran and McDonald (2011). Panel A of Table 4 reports the results using these bag of words indices. As shown in column (4), after controlling for financial and macroeconomic fundamentals, a one standard deviation increase in China related news is associated with a 0.93% increase in quarterly investment flows into Chinese assets, which corresponds to a 3.72% increase on an annualized basis. These estimates are consistent with and comparable in magnitude to our baseline results, providing additional support for the robustness of our index construction.

6.2 Robustness Analysis: Cross-language FinBERT Analysis

Benson et al. (2025) show that language differences in news coverage can matter for asset prices and market outcomes. Hence, we restimate our baseline analysis replacing the English language narrative indices with indices constructed using local language FinBERT models in German, French, and Spanish applied to domestic newspapers. Panel (B) of Table 4 reports results from this analysis. We find that domestic language narratives about China are also strongly predictive of cross border investment flows. As shown in column (4), after controlling for macroeconomic and financial fundamentals as well as fund and quarter fixed effects, a one standard deviation increase in the domestic language sentiment index is associated with a 0.75% increase in quarterly flows to China, corresponding to an annualized increase of approximately 2.98%. The estimated effect is highly statistically significant and economically meaningful, with a magnitude that is slightly smaller than, but comparable to, the baseline English language results.[19]

[19] One explanation for the modestly lower coefficients in the cross language specification is that domestic language newspapers exhibit less variation in their coverage of China. German, French, and Spanish language outlets tend to be more domestically oriented and devote a smaller share of reporting to China relative to global English language media such as the *Wall Street Journal*, the *Financial Times*, or major United States national newspapers. As a result, fluctuations in China related narratives are more muted in local

The strong statistical significance and consistent signs across languages reinforce that the results are not driven by linguistic artifacts or the choice of media language. Instead, they point to a robust underlying mechanism through which media narratives, whether expressed in global English outlets or domestic language media, shape cross border capital flows to China.

6.3 Robustness Analysis: Extracting Media Narrative Shocks

In the previous section, we show that media narratives about China from investors' domicile countries influence cross-border fund flows into Chinese assets, even after controlling for a range of macroeconomic and financial variables. However, the constructed media indices may still reflect a combination of underlying fundamentals in China along with the narrative framing that shapes how these fundamentals are perceived and interpreted. To address this concern, in this section, we conduct robustness checks by isolating media narrative shocks that are orthogonal to macroeconomic and financial fundamentals.

To isolate the narrative component ($narrativeshocks_{c,t}$) from the sentiment and risk indices, we estimate the following equation:

$$Index_{c,t} = g(fundamentals_t) + \varepsilon_{c,t},$$

where $index_{c,t}$ is the sentiment index constructed for domicile country c, $g(\cdot)$ is a function of current and lagged Chinese fundamentals, and $\varepsilon_{c,t}$ captures the narrative shocks, i.e., deviations from what fundamentals alone would predict.

In particular, we use a recursive vector autoregression (VAR) with Cholesky decompo-

language media markets, which mechanically limits the variation captured by the sentiment indices. This structural feature implies that multilingual sentiment indices display lower cross country dispersion, even though their informational content remains highly relevant. When the underlying variance of narratives is smaller, the estimated elasticity of capital flows is naturally attenuated, despite the relationship being precisely estimated and statistically strong. Moreover, international investors, particularly those allocating to emerging markets, rely heavily on English language financial news when forming global investment views. Consequently, English based indices capture a broader and more globally salient information set, which helps explain their somewhat larger estimated effects.

sition to identify $\varepsilon_{c,t}$. The baseline VAR specification is:

$$\boldsymbol{A}\boldsymbol{Y}_{\boldsymbol{t}} = \sum_{j=1}^{p} \boldsymbol{C}_{\boldsymbol{j}} \boldsymbol{Y}_{\boldsymbol{t-j}} + \boldsymbol{\varepsilon}_{\boldsymbol{t}}$$

where $\boldsymbol{A}$ and $\boldsymbol{C}_{\boldsymbol{j}}$ are $K \times K$ matrices of parameters, ε_t is a $K \times 1$ vector of innovation with $\varepsilon_t \sim N(0, \boldsymbol{\sigma})$ and $E[\varepsilon_t \varepsilon_s^{'}] = \boldsymbol{O}_K$ for all $s \neq t$. For the variables, the Cholesky restrictions result in the exclusion restriction on contemporaneous response in the matrix $\boldsymbol{A}$ to fit a just-identified model, with $\boldsymbol{A}$ to be a lower triangular matrix.

We include three lags ($p = 3$) based on the combined results of lag-order selection statistics, including final prediction error (FPE), Akaike's information criterion (AIC), Schwarz's Bayesian information criterion (BIC), and the Hannan and Quinn information criterion (HQIC).

The vector $\boldsymbol{Y}_{\boldsymbol{t}}$ includes the following variables, all in log form unless otherwise noted: the constructed media sentiment index; the stock return based on the Shanghai–Shenzhen CSI 300 index (Ret_t), capturing overall Chinese equity market performance; the volatility index (Vol_t), defined as the monthly volatility of the CSI 300; China's RMB effective exchange rate index (EX_t); the monthly year-over-year Consumer Price Index (CPI_t), reflecting inflation dynamics; and the seasonally adjusted industrial production index (IP_t), representing real economic activity.[20]

We order the variables in $\boldsymbol{Y}_{\boldsymbol{t}}$ based on standard macro-financial assumptions. (1) The media index is placed first, under the assumption that narrative shocks can contemporaneously affect all other variables, but are not themselves contemporaneously affected by them. (2) The stock return follows, as stock returns react immediately to news and are considered highly responsive. (3) The volatility index is ordered next, as it typically reacts to returns but exhibits more persistence and slower transmission to other variables. (4) The

[20]The media narrative index and stock returns are not log-transformed because both series can take negative values. The stock return is already expressed as the log difference of prices, and the sentiment index is centered around zero; therefore, taking logs is neither meaningful nor feasible.

exchange rate index is placed fourth, reflecting its sensitivity to financial variables while assuming it does not contemporaneously influence real economic activities. (5) The CPI is then included, based on its slower-moving nature and limited immediate responsiveness to financial shocks. (6) Finally, the industrial production index appears last, as it is assumed to be the most sluggish among the variables and does not respond contemporaneously to any of the preceding shocks.[21]

The first element of ε_t, corresponding to the media index, is interpreted as the media narrative shock. We estimate this structural VAR separately for each country and for both sentiment and risk indices to generate country-level media narrative shock series.

We repeat the baseline analysis using media narrative shocks identified from the recursive VAR model rather than the original narrative indices. Panel C of Table 4 reports the results. As shown in column (4), a one standard deviation increase in the media sentiment shock about China is associated with a 1.47% increase in quarterly investment flows into Chinese assets, which corresponds to an annualized increase of 5.88%. Consistent with the findings from the cross-country narrative decomposition, these results suggest that media narratives influence cross-border capital flows not only through fundamentals reflected in the indices, but also through the framing and interpretation of those fundamentals. This evidence supports a narrative-based transmission mechanism in which deviations in media sentiment that are orthogonal to macroeconomic and financial conditions have independent effects on international portfolio allocation.

6.4 Event Study Analysis: Arab Spring as a Global Narrative Shock

The baseline results document a strong association between domestic media sentiment toward China and cross-border investment flows. A key remaining question is whether this relationship reflects differences in how common global information is interpreted through domestic media narratives, rather than unobserved country-specific exposure to China.

[21] Our results are robust to alternative ordering schemes, including placing the stock return index first, followed by the media and volatility indices.

This section addresses this question by exploiting a plausibly exogenous global geopolitical shock to study how media narratives and capital flows respond across countries.

The Arab Spring triggered a worldwide reassessment of political stability, governance, and regime durability, elevating the salience of geopolitical risk in media narratives well beyond the countries directly affected. Although China's economic fundamentals were unchanged, this shift in global political narratives led media outlets to reinterpret China through a more geopolitical and risk-oriented lens.[22] We examine whether this common global shock generated differential changes in media narratives about China across investor countries and whether such narrative shifts were accompanied by changes in cross-border investment.

Our identification leverages cross-country heterogeneity in pre-existing exposure to geopolitical news. The key idea is that countries whose media environments are structurally more oriented toward geopolitical reporting may respond to the same global shock by re-framing narratives about China along geopolitical dimensions, rather than by reallocating attention toward China. This approach allows us to isolate differential narrative responses to a common event while holding constant the underlying information set. Consistent with this mechanism, we show that following the Arab Spring, countries with greater exposure to geopolitical reporting experience a pronounced deterioration in sentiment toward China, despite no corresponding differential change in media attention. These narrative shifts are accompanied by relative declines in portfolio flows to Chinese assets, supporting a narrative-based transmission mechanism in which global shocks affect cross-border investment through country-specific framing and interpretation rather than changes in information availability or China's underlying fundamentals.

[22] For example, *The Wall Street Journal* has reported on December 14, 2011, in the article "Group Urges Countries To Protect Web Speech" that "*While the Arab Spring, Occupy Wall Street and other movements have shown the potential of the Internet for organizing political protest, there has also been a backlash, with a number of governments stepping up their efforts to crack down on free speech in the digital sphere. China, which has long blocked access to Web sites deemed to be undesirable, said recently that it would step up monitoring of social media, messaging services and other forums in an effort to crack down on the publishing of 'harmful information.' ...*".

6.4.1 Measuring Geopolitical News Exposure

We construct a country-level measure of exposure to geopolitical news based on newspaper coverage. For each newspaper in our dataset, we identify geopolitical-related articles using the keyword list developed by Caldara and Iacoviello (2022). We then compute, for each newspaper, the share of geopolitical-related articles as a percentage of all articles published over the full sample period from 2007 to 2022. This measure captures persistent differences in editorial focus toward geopolitical issues rather than short-run responses to specific events.

We aggregate this measure to the country level by averaging across newspapers within each country. Countries are then ranked according to their average geopolitical news share, and we define an indicator

$$HighGeoPolitics_c = \mathbb{1}\{\text{GeoExposure}_c > \text{median}\},$$

which equals one for countries with above-median geopolitical news exposure and zero otherwise. Importantly, this classification is time-invariant and reflects long-run media orientation rather than responses to the Arab Spring itself.

6.4.2 Media Narrative Response to the Arab Spring

To examine how media narratives about China respond to the Arab Spring, we define a post-event indicator $Post_t$ equal to one for all months from January 2011 onward. We estimate the following difference-in-differences specification at the country–month level:

$$index_{c,t} = \alpha + \beta\,(Post_t \times HighGeoPolitics_c) + \gamma_c + \delta_t + \varepsilon_{c,t}, \tag{12}$$

where $index_{c,t}$ denotes alternative country-level media narrative indices about China, including a China-related attention index and an aggregate sentiment index. Country

fixed effects γ_c absorb time-invariant differences in media systems and baseline attitudes toward China, while month fixed effects δ_t capture global shocks common to all countries.

Panel A of Table 5 reports the results from this first-stage analysis across different event windows. We find no statistically significant differential change in media attention to China following the Arab Spring for countries with high geopolitical news exposure. In contrast, we observe a pronounced and persistent decline in aggregate sentiment toward China among these countries. The sentiment effect emerges gradually and becomes statistically significant over longer horizons, consistent with a re-framing of narratives rather than a mechanical reallocation of media coverage.

6.4.3 Capital Flow Response

We next examine whether these differential narrative responses are associated with subsequent changes in cross-border investment. Using fund-level data, we estimate:

$$Flow_{ic,t} = \theta\,(Post_t \times HighGeoPolitics_c) + \alpha_i + \alpha_t + X'_{c,t-1}\Phi + \varepsilon_{ic,t}, \tag{13}$$

where $Flow_{ic,t}$ measures fund i's capital flow to Chinese assets in quarter t. The specification includes fund fixed effects α_i, which absorb time-invariant investment mandates and fund characteristics, and quarter fixed effects α_t, which capture global investment conditions. The vector $X_{c,t}$ includes a set of macroeconomic and financial controls same as the baseline specification.

Results reported in Panel B of Table 5 suggest that, following the Arab Spring, funds domiciled in countries with high geopolitical news exposure reduce their investments in Chinese assets significantly more than funds from countries with lower exposure. The divergence in capital flows becomes economically and statistically significant over medium- to long-run horizons, even after controlling for macroeconomic fundamentals and fund-specific factors.

These results suggest that the Arab Spring triggered a global shift in geopolitical narratives that disproportionately affected countries whose media environments are more oriented toward geopolitical reporting. While media attention to China does not change differentially across countries, sentiment toward China deteriorates significantly in high-geopolitical-exposure countries, and this deterioration is accompanied by a relative decline in institutional investment flows into Chinese assets. These findings are consistent with a narrative-based transmission mechanism in which global geopolitical shocks reshape the tone and interpretation of media coverage about China, influencing investor perceptions and portfolio allocation decisions independently of China's underlying economic fundamentals.

7 Identifying the Media Narrative Impacts: A Shift–Share Method Using Non-China News

To further address the concern that media sentiment about China may simply mirror changes in China's macroeconomic or financial fundamentals, this section employs a shift–share identification strategy to isolate the causal impact of media narratives. We construct narrative shocks using topic-level sentiment from non-China news coverage in the *Wall Street Journal* and interact these shocks with predetermined country-level topic composition. Because the shocks are external to China and common across investor countries, the identifying variation reflects differential exposure to global narrative shifts rather than reactions to contemporaneous China-specific economic developments.

7.1 Shift–Share (Bartik) Instrument Based on Topic Composition

We first develop a shift–share (Bartik) identification strategy to estimate the causal effect of media narratives about China on cross-border portfolio flows. The key idea is to instrument investor-country media sentiment using predetermined differences in topic

composition interacted with exogenous, time-varying shifts in the global narrative environment. By construction, this approach isolates variation in domestic media sentiment that is driven by global narrative developments unrelated to China itself, rather than by unobserved investor-country determinants of capital flows or contemporaneous China-specific fundamental shocks.

Predetermined Topic Composition. In the previous section, we have shown that the aggregate sentiment index can be equivalently constructed as a topic-share-weighted average of within-topic sentiment,

$$sen_{c,t} = \sum_{k=1}^{K} w_{c,k,t}\, sen_{c,t}^{(k)}, \tag{14}$$

where the topic weights $w_{c,k,t}$ are measured as the fraction of China-related articles in topic k among all China-related articles published in country c at time t.To ensure that topic composition is predetermined with respect to subsequent capital flow shocks, we fix country–topic weights using a pre-sample period $\mathcal{T}_0$ spanning January 2007 to December 2011. This period predates major China-specific financial and geopolitical regime shifts, so topic shares primarily reflect persistent editorial attention rather than responses to later capital-flow dynamics. Specifically, we define

$$\bar{w}_{c,k} = \frac{1}{|\mathcal{T}_0|} \sum_{t \in \mathcal{T}_0} w_{c,k,t}, \qquad \sum_{k=1}^{K} \bar{w}_{c,k} = 1. \tag{15}$$

These predetermined shares capture persistent cross-country differences in the attention structure of China-related media coverage across topics and are held fixed throughout the estimation period.

Global Topic-Level Narrative Shocks. We measure global narrative shocks using sentiment indices constructed from *non-China* news coverage in the *Wall Street Journal* over the

period 2007–2022, using the same topic classification and sentiment construction procedures described in Section 2. By construction, these shocks capture shifts in the global narrative environment. such as changes in global risk sentiment, macroeconomic outlook, or geopolitical tension, that are plausibly exogenous to China-specific developments and investor-country portfolio allocation decisions vis-à-vis China.

Construction of the Shift–Share Instrument. The shift–share instrument for country c at time t is defined as the exposure of that country's predetermined topic composition to global narrative shocks,

$$Z_{c,t} = \sum_{k=1}^{K} \bar{w}_{c,k} \, G_{k,t}. \tag{16}$$

The instrument varies over time with global topic-level narrative shocks and varies across countries with persistent differences in topic attention. Country fixed effects absorb all time-invariant differences in narrative stance, while time fixed effects remove common global shocks, so identification relies on differential exposure to the same global narrative shifts.

7.2 Methodology

First Stage. At the first stage, we assess the relevance of the instrument using the first-stage regression

$$sen_{c,t} = \pi Z_{c,t} + \alpha_c + \alpha_t + X'_{c,t}\delta + u_{c,t} \tag{17}$$

where α_c and α_t denote country and month fixed effects, respectively. A positive and statistically significant π indicates that global narrative shocks, interacted with predetermined topic composition, generate systematic variation in domestic media sentiment about China. The fitted monthly values from this regression are saved and subsequently standardized and aggregated to the quarterly frequency, producing the quarterly instrumented media sentiment index $\widehat{sen}^{IV}_{c,t}$ used in the second-stage.

Second Stage. The causal effect of media sentiment on portfolio flows is estimated using

$$Flow_{ic,t} = \beta_0 + \beta_1 \widehat{sen}^{IV}_{c,t-1} + \beta_2 num_{c,t-1} + X'_{c,t-1}\Phi + \alpha_i + \alpha_t + \varepsilon_{ic,t} \tag{18}$$

The coefficient β_1 captures the impact of investor-country media sentiment about China on capital flows into Chinese assets.

Identification Assumptions. This ientification relies on two assumptions. First, predetermined topic shares $\bar{w}_{c,k}$ are orthogonal to future innovations in portfolio flows, ensuring that media attention structure is not chosen in response to subsequent shocks.[23] Second, global narrative shocks $G_{k,t}$ constructed from non-China WSJ coverage are orthogonal to investor-country-specific shocks to flows into China. These conditions imply that the instrument affects portfolio flows only through its impact on domestic media sentiment about China.

7.3 Identifying the Media Narrative Impacts

Table 6 reports results from the shift–share identification strategy that uses global narrative shocks from non-China *Wall Street Journal* coverage to instrument investor-country media sentiment about China. Panel A presents the second-stage estimates of the effect of instrumented media sentiment on capital flows into Chinese assets. Across all specifications, the coefficient on lagged instrumented sentiment, $\widehat{sen}^{IV}_{c,t-1}$, is positive, economically meaningful, and statistically significant, indicating an independent narrative channel that operates beyond standard macroeconomic and financial factors. Panel B reports the corresponding first-stage results at the monthly frequency. The shift–share instrument constructed from predetermined topic composition and non-China WSJ narrative shocks

[23] To address the concern that topic composition may proxy for broader economic exposure to China, the second-stage specification augments the baseline regression with interactions between observable China-specific shocks and time-invariant country exposure measures, such as trade intensity with China. This ensures that identification does not rely on differential economic exposure to China.

strongly predicts domestic media sentiment about China, with a first-stage coefficient of 0.90 and a robust F-statistic of 16.26. This confirms that global narrative shocks, when filtered through persistent cross-country differences in topic attention, generate substantial and systematic variation in domestic media sentiment. Taken together, these results provide consistent evidence that exogenous shifts in the global narrative environment translate into changes in investor-country media sentiment and, in turn, have a causal effect on cross-border capital flows into China.

8 Asymmetric and Higher-Moment Effects of Media Narratives

Asymmetric Effects of Media Narratives. A large literature documents that news media disproportionately emphasize negative events, exhibiting a systematic bias toward unfavorable reporting (Goidel and Langley, 1995; Damstra and Boukes, 2021; van Binsbergen, Bryzgalova, Mukhopadhyay and Sharma, 2024). At the same time, extensive evidence suggests that individuals tend to react more strongly to negative than to positive information (Holbrook, Krosnick, Visser, Gardner and Cacioppo, 2001; Soroka, 2006). Motivated by these findings, we examine whether cross-border institutional investors respond asymmetrically to positive versus negative media narratives about China. In particular, we ask whether negative narratives exert a stronger influence on portfolio allocation decisions, or whether positive narratives play a more salient role in shaping investment flows given prevailing pessimistic priors about emerging markets.

To distinguish between positive and negative news, we construct two separate indices based on word frequencies.[24] Using the positive and negative word lists from Loughran and McDonald (2011), the positive media narrative index, $pos_{c,t}$, is computed as the frequency of positive words in each article scaled by article length, averaged across all articles within a given media outlet, and then averaged across all outlets within each

[24]Similarly, Rey, Stavrakeva and Tang (2024) define two risk news indices, "risk on" and "risk off," to study the relationship between exchange rates and the global network of equity holdings.

country. The negative media narrative index, $neg_{c,t}$, is constructed analogously using negative word frequencies.[25]

Panel A of Table 7 reports the results from specifications that jointly include the positive and negative media narrative indices. Consistent with conventional predictions based on negativity bias, we find that negative narratives exert a statistically significant influence on capital flows, whereas positive narratives do not. As shown in column (4), a one standard deviation increase in the negative sentiment index is associated with a 0.69% decline in quarterly fund flows into Chinese assets, while the estimated effect of the positive sentiment index is small and statistically insignificant. These results indicate that adverse media narratives play a dominant role in shaping cross-border institutional investment decisions.

This asymmetry is consistent with the idea that negative news carries greater salience in investors' information processing, particularly in cross-border settings characterized by uncertainty and limited transparency. In an environment where media coverage already exhibits a structural bias toward adverse events, negative narratives may amplify downside concerns and trigger portfolio reallocation away from Chinese assets. By contrast, positive narratives appear insufficient to offset entrenched negative perceptions, suggesting that downside information is more influential than upside signals in driving international portfolio flows.

Higher Moment Effects of Media Narratives. Hassan, Hollander, Van Lent and Tahoun (2019) argue that changes in sentiment reflect shifts in expected average outcomes, corresponding to the first moment of the distribution, whereas changes in risk capture variation or uncertainty around those outcomes, corresponding to the second moment. To examine higher moment effects of media narratives, we follow the approach in Hassan, Hollander, Van Lent and Tahoun (2019) and construct a risk index, denoted by $risk_{c,t}$ for investors

[25] As an alternative approach, we construct positive news indices using only articles with positive sentiment scores and negative news indices using only articles with negative sentiment scores. Our results are robust to this alternative construction.

domiciled in country c in month t. The detailed construction of this index is described in Appendix B. This risk index captures country specific perceptions of uncertainty about China as reflected in each country's media coverage.

As shown in Panel B of Table 7 column (8), after controlling for cross country macroeconomic and financial fundamentals, a one standard deviation increase in the risk perception index is associated with a 0.66% quarterly decline in fund flows into Chinese assets. This result indicates that heightened perceived risk conveyed by media in investors' domicile countries has a statistically and economically meaningful negative effect on cross border investment flows.

9 Conclusion

We study how media narratives about a common foreign economy differ across investor countries and how such differences shape international capital allocation. Using a large corpus of newspapers from multiple investor countries, we construct novel measures of media attention and sentiment toward China and document substantial, persistent cross-country dispersion in media sentiment, even when outlets cover the same underlying events. These findings demonstrate that narratives about a foreign economy are not global or uniform, but instead reflect country-specific framing and interpretation.

We unpack the sources of this narrative disagreement. By decomposing aggregate sentiment into topic attention and within-topic sentiment, we show that nearly all cross-country variation in sentiment is driven by differences in interpretation rather than differences in what countries choose to cover. Further decompositions reveal that narrative disagreement reflects both slow-moving country-specific belief levels and heterogeneous responses to new information, with their relative importance varying across topics. These results provide direct evidence that investors in different countries interpret the same foreign economy through distinct narrative lenses.

We then link narrative disagreement to behavior by showing that domestic media senti-

ment toward China has an effect on cross-border portfolio flows, even after controlling for fundamentals and global financial conditions. This relationship is identified using exogenous global narrative shocks from non-China news coverage, which generate differential changes in domestic media sentiment across investor countries. This evidence highlights domestic media narratives as a distinct channel shaping belief formation and international portfolio choice, complementing existing work that treats information frictions as latent or proxied by distance, institutions, or familiarity.

Our results open several avenues for future research. First, media-based measures of narrative disagreement could be used to study cross-country differences in reactions to global shocks, policy announcements, or geopolitical events. Second, extending the analysis to other destination economies would shed light on whether narrative disagreement is more pronounced in environments characterized by opacity or policy uncertainty. Finally, linking media narratives to asset prices and risk premia across borders would further illuminate the role of belief formation in global financial markets. Overall, this paper shows that understanding international capital flows requires not only measuring fundamentals, but also accounting for the narratives through which investors interpret them.

References

Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz (2006) "Learning and disagreement in an uncertain world," NBER Working Paper 12648, National Bureau of Economic Research.

Agarwal, Isha, Grace Weishi Gu, and Eswar Prasad (2020) "The determinants of China's international portfolio equity allocations," *IMF Economic Review*, 68, 643–692.

Ahir, Hites, Nicholas Bloom, and Davide Furceri (2022) "The world uncertainty index," Working Paper 29763, National Bureau of Economic Research.

Ahmed, Shaghil and Andrei Zlate (2014) "Capital flows to emerging market economies: A brave new world?" *Journal of International Money and Finance*, 48, 221–248.

Alok, Shashwat, Apoorva Javadekar, Nitin Kumar, and Russ Wermers (2022) "Economic policy uncertainty and global portfolio allocations," *Available at SSRN 4193637*.

Andrade, Sandro C and Vidhi Chhaochharia (2010) "Information immobility and foreign portfolio investment," *The Review of Financial Studies*, 23 (6), 2429–2463.

Arteaga-Garavito, Maria Jose, Ric Colacito, Mariano (Max) Massimiliano Croce, and Biao Yang (2024) "International climate news," *Working Paper*.

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.

Baker, Scott, Nicholas Bloom, Steven J Davis, and Marco C Sammon (2021) "What triggers stock market jumps?" Working Paper 28687, National Bureau of Economic Research.

Baker, Scott R, Nicholas Bloom, and Steven J Davis (2016) "Measuring economic policy uncertainty," *The Quarterly Journal of Economics*, 131 (4), 1593–1636.

Benson, Kevin, Ing-Haw Cheng, Maurice Granger et al. (2025) "Measuring Price Effects of Multilingual Global News with Large Language Models."

Besley, Timothy and Andrea Prat (2006) "Handcuffs for the grabbing hand? Media capture and government accountability," *American economic review*, 96 (3), 720–736.

van Binsbergen, Jules H, Svetlana Bryzgalova, Mayukh Mukhopadhyay, and Varun Sharma (2024) "(Almost) 200 years of news-based economic sentiment," Working Paper 32026, National Bureau of Economic Research.

Bray, Margaret M. and N. E. Savin (1986) "Rational expectation equilibrium, learning, and model specification," *Econometrica*, 54 (5), 1129–1160.

Bybee, Leland, Bryan Kelly, Asaf Manela, and Dacheng Xiu (2024) "Business news and business cycles," *The Journal of Finance*, 79 (5), 3105–3147.

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.

Chahrour, Ryan, Kristoffer Nimark, and Stefan Pitschner (2021) "Sectoral media focus and aggregate fluctuations," *American Economic Review*, 111 (12), 3872–3922.

Chan, Kalok, Vicentiu Covrig, and Lilian Ng (2005) "What determines the domestic bias and foreign bias? Evidence from mutual fund equity allocations worldwide," *The Journal of Finance*, 60 (3), 1495–1534.

Conte, Maddalena, Pierre Cotterlaz, Thierry Mayer et al. (2022) *The CEPII gravity database*, 5: CEPII Paris, France.

Cookson, J Anthony, Chukwuma Dim, and Marina Niessner (2024) "Disagreement on the Horizon," *Available at SSRN*.

Cookson, J Anthony, Joseph E Engelberg, and William Mullins (2020) "Does partisanship shape investor beliefs? Evidence from the COVID-19 pandemic," *The Review of Asset Pricing Studies*, 10 (4), 863–893.

——— (2023) "Echo chambers," *The Review of Financial Studies*, 36 (2), 450–500.

Cookson, J Anthony and Marina Niessner (2020) "Why don't we agree? Evidence from a social network of investors," *The Journal of Finance*, 75 (1), 173–228.

Coppola, Antonio, Matteo Maggiori, Brent Neiman, and Jesse Schreger (2021) "Redrawing the map of global capital flows: The role of cross-border fnancing and tax havens," *The Quarterly Journal of Economics*, 136 (3), 1499–1556.

Cyert, Richard M. and Morris H. DeGroot (1974) "Rational expectations and Bayesian analysis," *Journal of Political Economy*, 82 (3), 521–536.

Damstra, Alyt and Mark Boukes (2021) "The economy, the news, and the public: A longitudinal study of the impact of economic news on economic evaluations and expectations," *Communication Research*, 48 (1), 26–50.

Engelberg, Joseph and Christopher A. Parsons (2011) "The Causal Impact of Media in Financial Markets," *The Journal of Finance*, 66 (1), 67–97.

Flynn, Joel P and Karthik Sastry (2024) "The macroeconomics of narratives," Working Paper 32602, National Bureau of Economic Research.

Forbes, Kristin J and Francis E Warnock (2012) "Capital flow waves: Surges, stops, flight, and retrenchment," *Journal of International Economics*, 88 (2), 235–251.

Froot, Kenneth A and Tarun Ramadorai (2005) "Currency returns, intrinsic value, and institutional-investor flows," *The Journal of Finance*, 60 (3), 1535–1566.

Gentzkow, Matthew and Jesse M Shapiro (2006) "Media bias and reputation," *Journal of political Economy*, 114 (2), 280–316.

——— (2010) "What drives media slant? Evidence from US daily newspapers," *Econometrica*, 78 (1), 35–71.

Gentzkow, Matthew, Jesse M Shapiro, and Michael Sinkinson (2014) "Competition and ideological diversity: Historical evidence from us newspapers," *American Economic Review*, 104 (10), 3073–3114.

Ghosh, Atish R, Mahvash S Qureshi, Jun Il Kim, and Juan Zalduendo (2014) "Surges," *Journal of International Economics*, 92 (2), 266–285.

Goidel, Robert K and Ronald E Langley (1995) "Media coverage of the economy and aggregate economic evaluations: Uncovering evidence of indirect media effects," *Political Research Quarterly*, 48 (2), 313–328.

Gulen, Huseyin and Mihai Ion (2016) "Policy uncertainty and corporate investment," *The Review of Financial Studies*, 29 (3), 523–564.

Gurevich, Tamara (2018) "The dynamic gravity dataset: 1948–2016," *USITC Working Paper, 2018–02–A*.

Haltiwanger, John and Michael Waldman (1985) "Rational expectations and the limits of rationality: An analysis of heterogeneity," *American Economic Review*, 75 (3), 326–340.

Handley, Kyle and J Frank Li (2020) "Measuring the effects of firm uncertainty on economic activity: New evidence from one million documents," Working Paper 27896, National Bureau of Economic Research.

Harris, Milton and Artur Raviv (1993) "Differences of opinion make a horse race," *The Review of Financial Studies*, 6 (3), 473–506.

Hassan, Tarek A, Stephan Hollander, Laurence Van Lent, and Ahmed Tahoun (2019) "Firm-level political risk: Measurement and effects," *The Quarterly Journal of Economics*, 134 (4), 2135–2202.

Hassan, Tarek A, Jesse Schreger, Markus Schwedeler, and Ahmed Tahoun (2024) "Sources and transmission of country risk," *Review of Economic Studies*, 91 (4), 2307–2346.

Head, Keith and Thierry Mayer (2014) "Gravity equations: Workhorse, toolkit, and cookbook," in *Handbook of international economics*, 4, 131–195: Elsevier.

Hoberg, Gerard and Asaf Manela (2025) "The Natural Language of Finance," *Available at SSRN 5119322*.

Holbrook, Allyson L, Jon A Krosnick, Penny S Visser, Wendi L Gardner, and John T Cacioppo (2001) "Attitudes toward presidential candidates and political parties: Initial optimism, inertial first impressions, and a focus on flaws," *American Journal of Political Science*, 930–950.

Horn, Sebastian, Carmen M Reinhart, and Christoph Trebesch (2021) "China's overseas lending," *Journal of International Economics*, 133, 103539.

Hu, Allen (2024) "Financial news production," *Working Paper*.

Hutchison, Michael M and Ilan Noy (2006) "Sudden stops and the Mexican wave: Currency crises, capital flow reversals and output loss in emerging markets," *Journal of Development Economics*, 79 (1), 225–248.

Jiang, Zhengyang (2025) "China's Overseas Lending in Global Finance Cycle,"Technical report, National Bureau of Economic Research.

Kandel, Eugene and Neil D Pearson (1995) "Differential interpretation of public signals and trade in speculative markets," *Journal of Political Economy*, 103 (4), 831–872.

Kandel, Eugene and Benjamin Z. Zilberfarb (1999) "Differential interpretation of information in inflation forecasts," *Review of Economics and Statistics*, 81 (2), 217–226.

Karolyi, G. Andrew, David T. Ng, and Eswar S. Prasad (2020) "The coming wave: where do emerging market investors put their money?" *Journal of Financial and Quantitative Analysis*, 55 (1), 1–46.

Lahiri, Kajal and Fushang Liu (2006) "Modelling multi-period inflation uncertainty using a panel of density forecasts," *Journal of Applied Econometrics*, 21 (8), 1199–1219.

Lahiri, Kajal and Xuguang Sheng (2008) "Evolution of forecast disagreement in a Bayesian learning model," *Journal of Econometrics*, 144 (2), 325–340.

Lee, Annie Soyean and Charles Engel (2024) "US liquid government liabilities and emerging market capital flows," Working Paper 32572, National Bureau of Economic Research.

Leuz, Christian, Karl V Lins, and Francis E Warnock (2009) "Do foreigners invest less in poorly governed firms?" *The Review of Financial Studies*, 22 (8), 3245–3285.

Loughran, Tim and Bill McDonald (2011) "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," *The Journal of Finance*, 66 (1), 35–65.

Maggiori, Matteo, Brent Neiman, and Jesse Schreger (2020) "International currencies and capital allocation," *Journal of Political Economy*, 128 (6), 2019–2066.

Manela, Asaf and Alan Moreira (2017) "News Implied Volatility and Disaster Concerns," *Journal of Financial Economics*, 123 (1), 137–162.

Mankiw, N Gregory and Ricardo Reis (2002) "Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve," *The Quarterly Journal of Economics*, 117 (4), 1295–1328.

Mullainathan, Sendhil and Andrei Shleifer (2005) "The market for news," *American economic review*, 95 (4), 1031–1053.

Nedopil, Christoph (2025) "Countries of the Belt and Road Initiative," *Shanghai, Green Finance & Development Center, FISF Fudan University, www.greenfdc.org*.

Nimark, Kristoffer P and Stefan Pitschner (2019) "News media and delegated information choice," *Journal of Economic Theory*, 181, 160–196.

Pesaran, M. Hashem (2006) "Estimation and inference in large heterogeneous panels with a multifactor error structure," *Econometrica*, 74 (4), 967–1012.

Portes, Richard and Helene Rey (2005) "The determinants of cross-border equity flows," *Journal of International Economics*, 65 (2), 269–296.

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

Rey, Helene, Vania Stavrakeva, and Jenny Tang (2024) "Currency centrality in equity markets, exchange rates and global financial cycles," Working Paper 33003, National Bureau of Economic Research.

Shiller, Robert J (2017) "Narrative economics," *American Economic Review*, 107 (4), 967–1004.

——— (2020) *Narrative economics: How stories go viral and drive major economic events*: Princeton University Press.

Shoemaker, Pamela J and Timothy Vos (2009) *Gatekeeping theory*: Routledge.

Sims, Christopher A (2003) "Implications of rational inattention," *Journal of monetary Economics*, 50 (3), 665–690.

Soroka, Stuart N (2006) "Good news and bad news: Asymmetric responses to economic information," *The Journal of Politics*, 68 (2), 372–385.

Tetlock, Paul C. (2007) "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *The Journal of Finance*, 62 (3), 1139–1168.

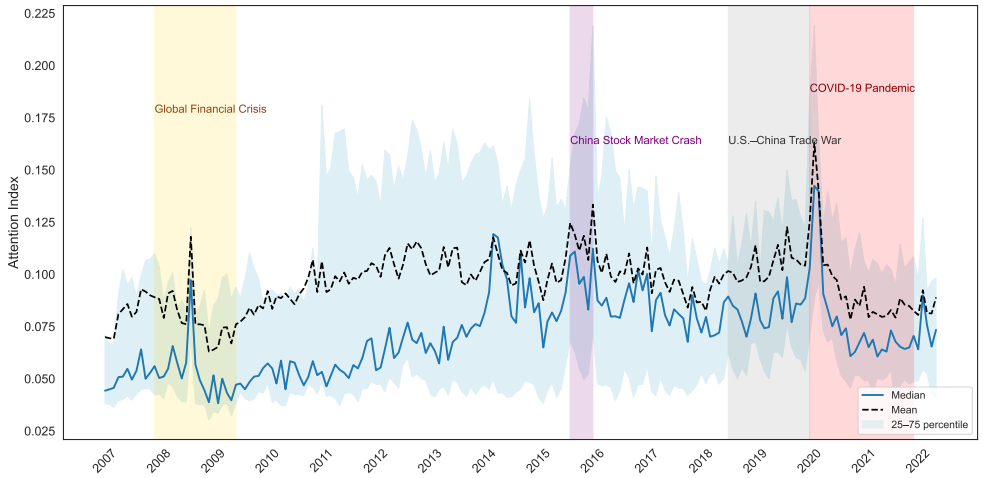
Townsend, Robert M. (1983) "Forecasting the forecasts of others," *Journal of Political Economy*, 91 (4), 546–588.

Van Nieuwerburgh, Stijn and Laura Veldkamp (2009) "Information immobility and the home bias puzzle," *The Journal of Finance*, 64 (3), 1187–1215.

Varian, Hal (1989) "Differences of opinion in financial markets," in Stone, C. ed. *Financial Risk: Theory, Evidence and Implications*, 3–37, Boston: Kluwer Academic Publishers.

Figures and Tables

(a) Attention index



(b) Sentiment index

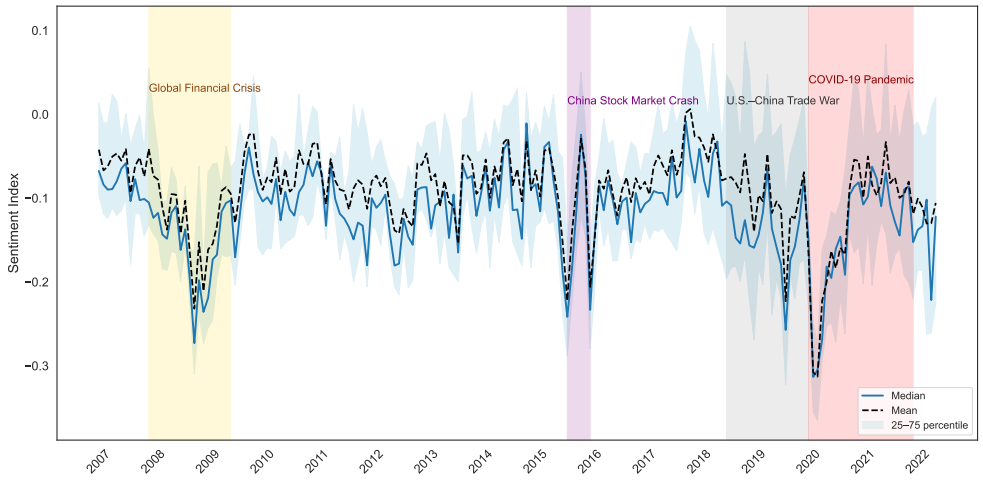
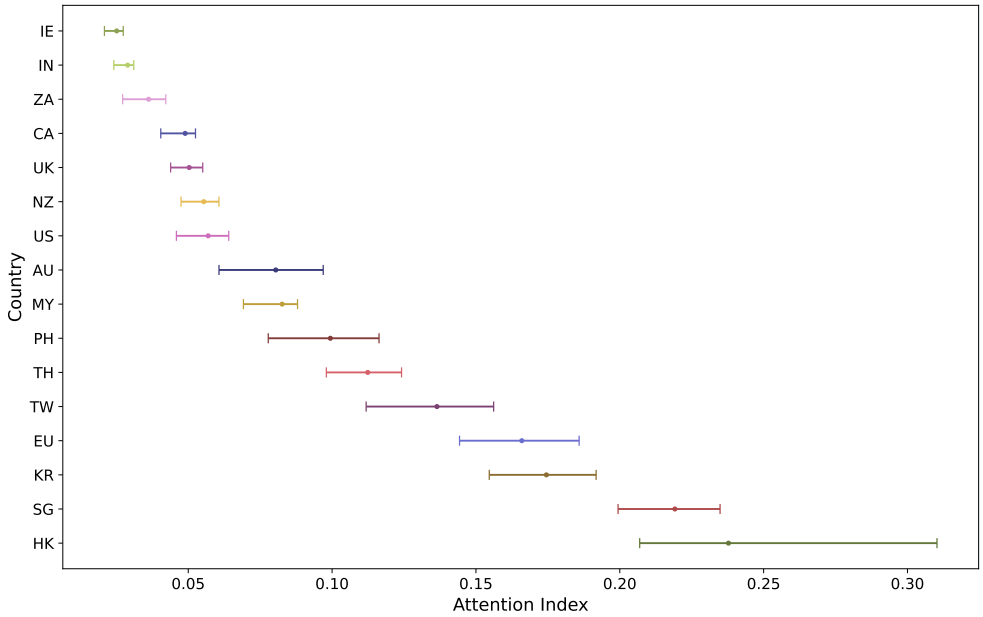


Figure 1: China Attention and Sentiment Indices

Notes: This figure plots the cross-country distribution of (a) the attention index and (b) the sentiment index for China over the period 2007–2022, constructed using the methodology described in Section 2. The black dashed line shows the cross-country mean, the solid blue line shows the median, and the shaded blue band represents the interquartile range (25th–75th percentiles).

(a) Attention index



(b) Sentiment index

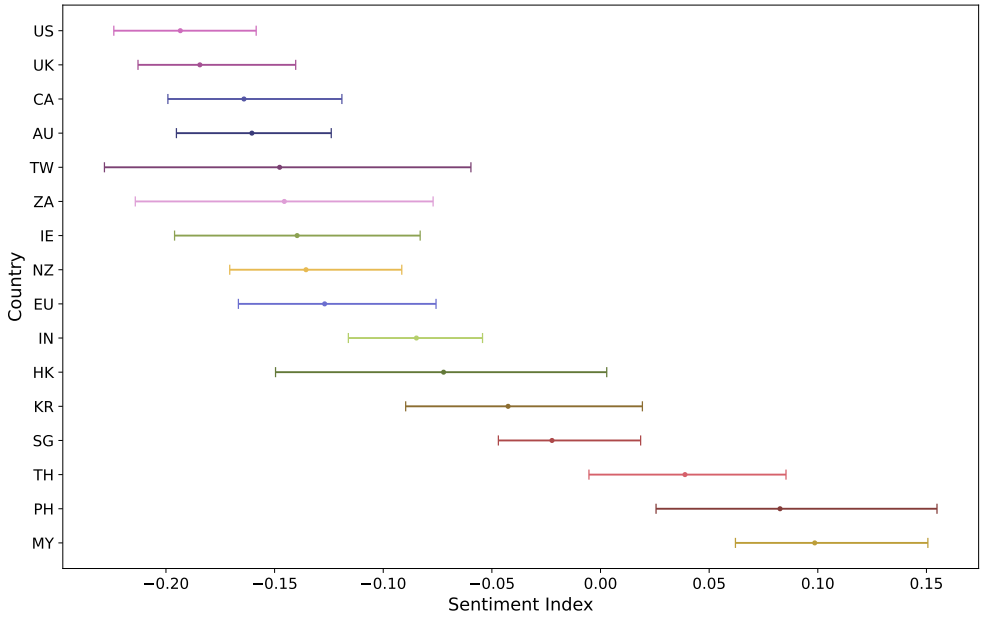
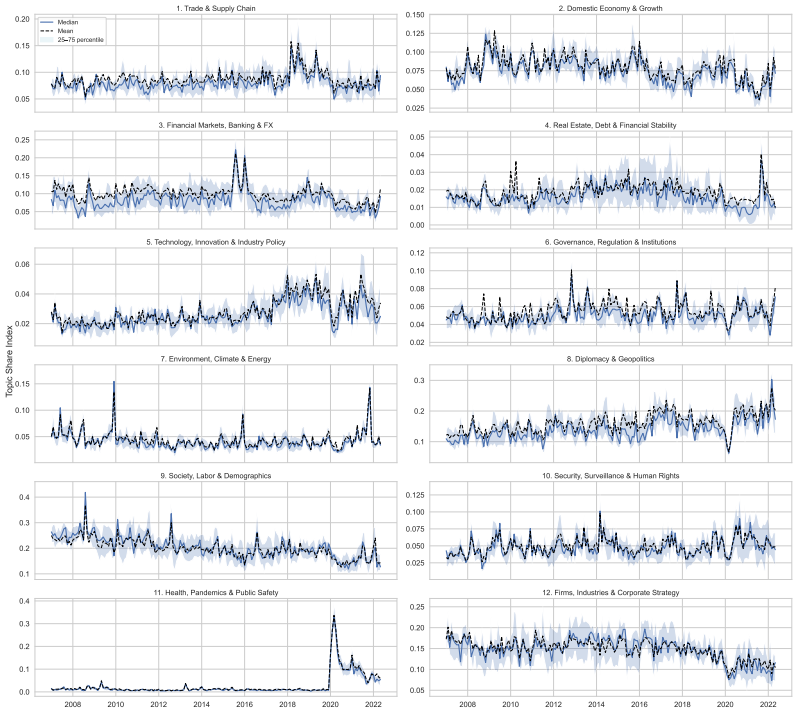


Figure 2: Cross-country 50th percentile ranges by index type

Notes: This figure illustrates the average cross-country dispersion in media narratives about China over the period 2007–2022. For each country in the sample, we plot the 50-th percentile range, defined as the difference between the 25th and 75th percentiles, of (a) the China-related attention index and (b) the China-related sentiment index, averaged over the sample period. All indices are constructed using the methodology described in Section 2.

(a) Topic share index



(b) Topic sentiment index

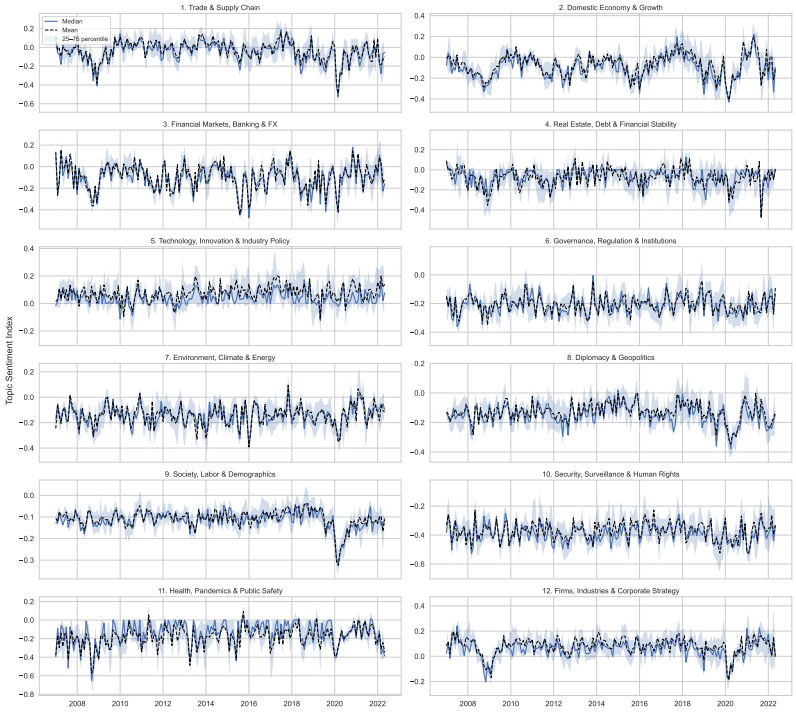
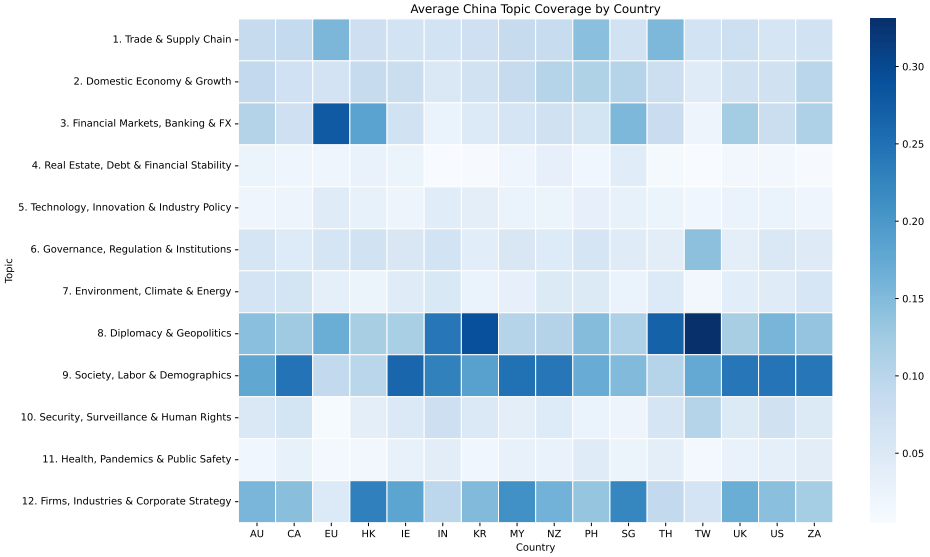


Figure 3: Topic Share and Sentiment Indices

Notes: This figure plots the cross-country distribution of (a) the topic share index and (b) the topic sentiment index for China over the period 2007–2022, constructed using the methodology described in Section 2. The black dashed line shows the cross-country mean, the solid blue line shows the median, and the shaded blue band represents the interquartile range (25th–75th percentiles).

(a) Average Topic Share Index



(b) Average Topic Sentiment Index

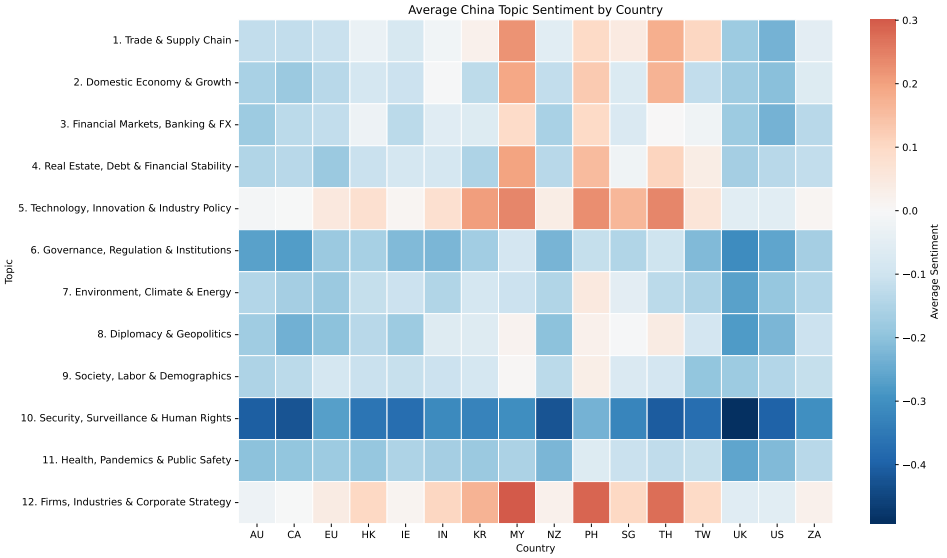


Figure 4: Average Topic Share and Sentiment by Country

Notes: This figure plots (a) the average topic share index and (b) the average topic sentiment index by country (horizontal axis) and topic (vertical axis) over the full sample period, constructed using the methodology described in Section 2. In panel (a), darker shading indicates a higher relative share of reporting devoted to each topic within all China-related news articles. In panel (b), red shading reflects more positive sentiment and blue shading reflects more negative sentiment, with darker colors indicating stronger sentiment intensity in either direction.

1. Trade & Supply Chain



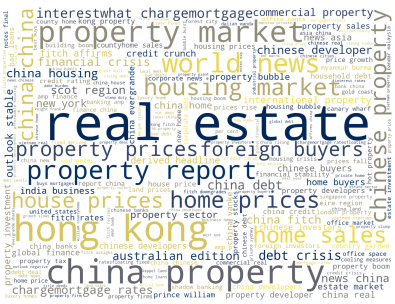
2. Domestic Economy & Growth



3. Financial Markets, Banking & FX



4. Real Estate, Debt & Financial Stability



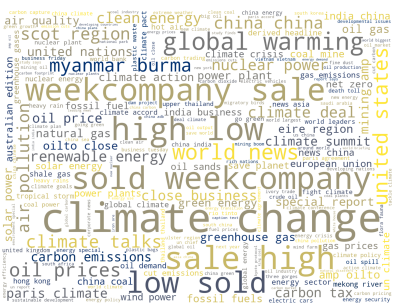
5. Tech, Innovation & Industry Policy



6. Governance, Regulation & Institutions



7. Environment, Climate & Energy



8. Diplomacy & Geopolitics



9. Society, Labor & Demographics



10. Security, Surveillance & Human Rights



11. Health, Pandemics & Public Safety



12. Firms, Industries & Corporate Strategy



Figure 5: Bigram Word Clouds Across Topics

Notes: This figure presents the bigram word clouds for the 12 topics used in our analysis.

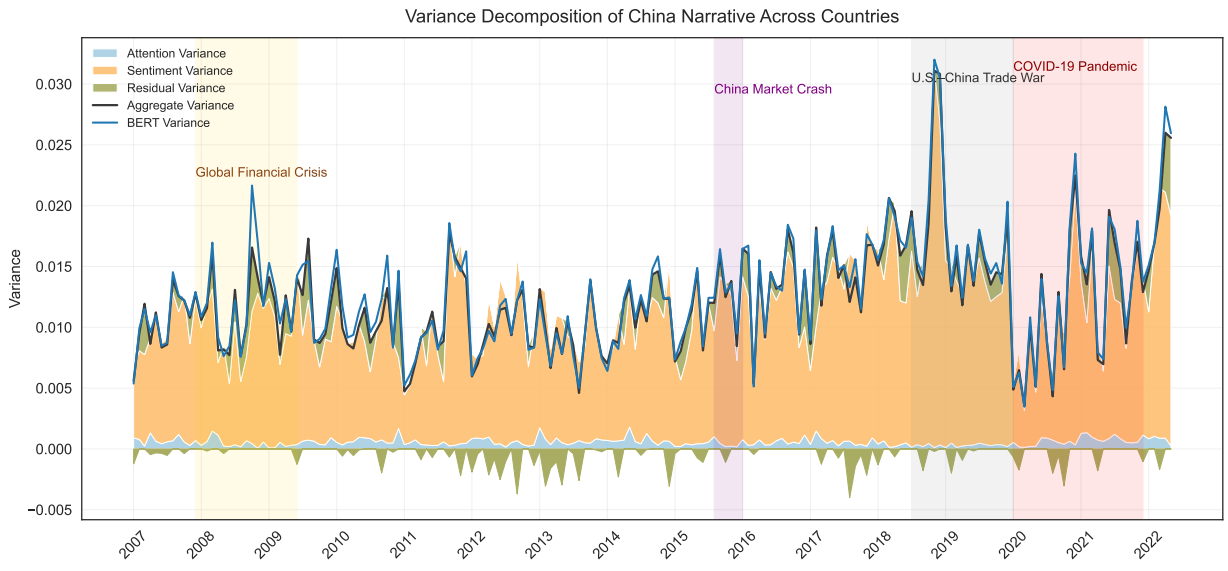
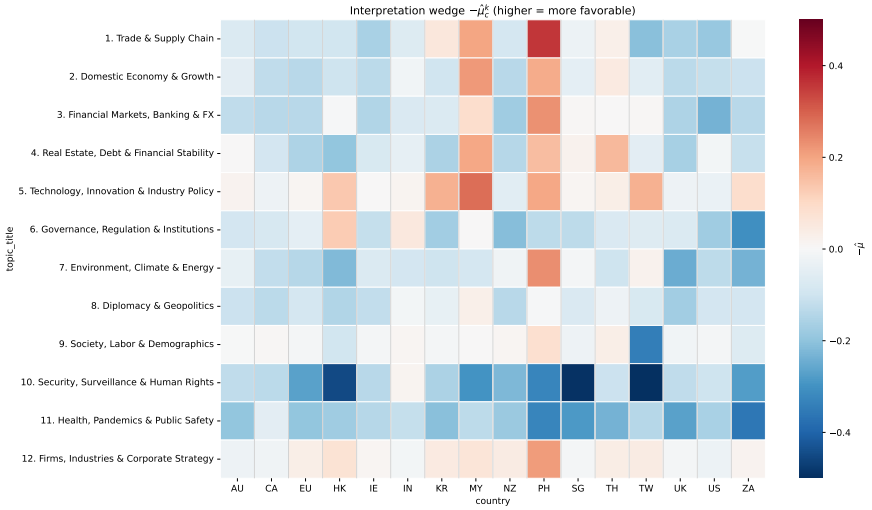


Figure 6: Variance Decomposition of China Narratives Across Countries

Notes: This figure reports a variance decomposition of cross-country dispersion in aggregate media sentiment toward China. For each month, aggregate sentiment is constructed as a weighted average of topic-level sentiment, where the weights reflect the allocation of media attention across topics within each country. The solid lines show the total cross-country variance of the aggregate sentiment index constructed from the decomposed measure (black) and from the raw index (blue). The shaded areas decompose this variance into three components. The attention component captures variation arising from cross-country differences in topic coverage, holding within-topic sentiment fixed at its cross-country average. The sentiment component captures variation arising from cross-country differences in sentiment conditional on topic, holding topic attention fixed at its cross-country average. The residual component reflects the interaction between topic attention and within-topic sentiment. All variances are computed across countries at a monthly frequency over the period 2007–2022. A detailed description of the methodology is provided in Section 4.1.

(a) Interpretation wedge $-\widehat{\mu}_c^k$ (higher = more favorable interpretation)



(b) Updating intensity $\widehat{\kappa}_c^k = 1 - \widehat{\lambda}_c^k$ (weight on new information)

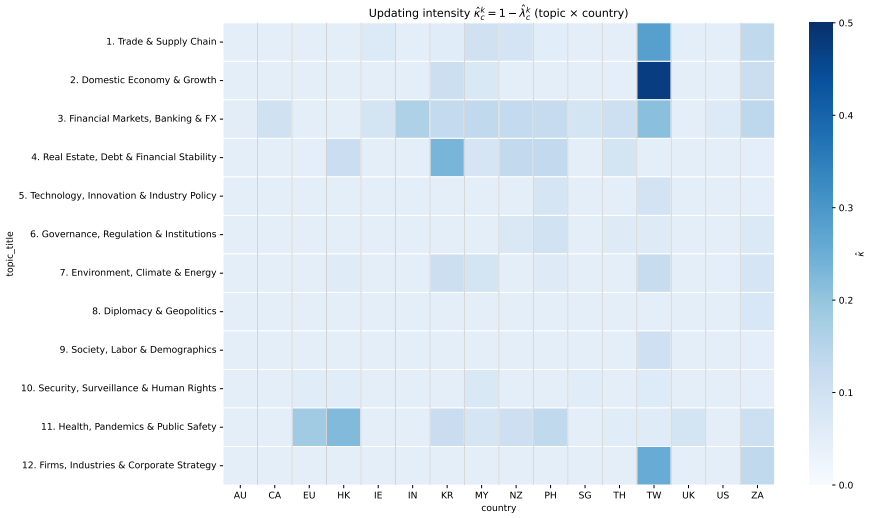
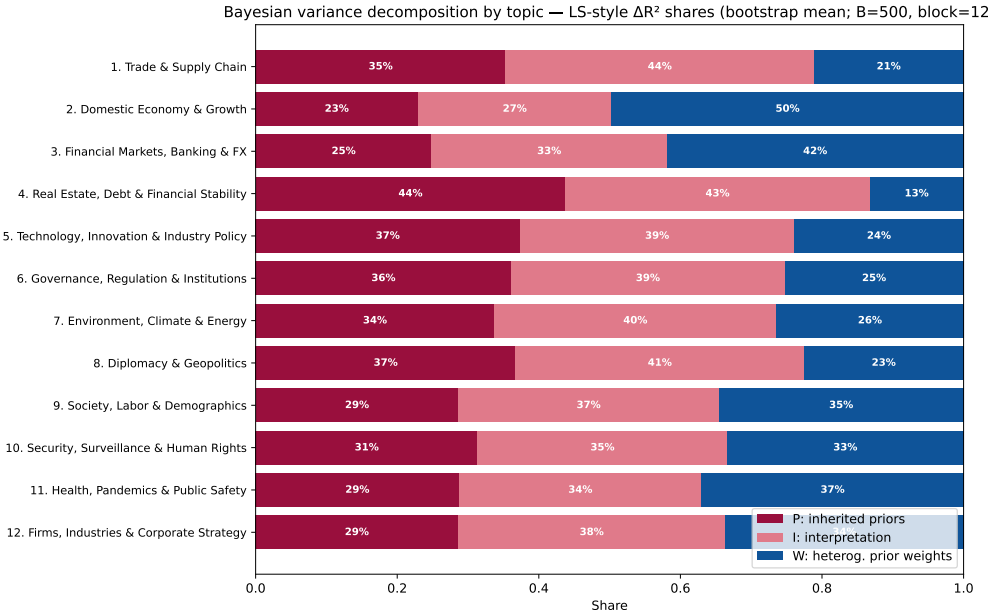


Figure 7: Country-level interpretation and updating

Notes: Panel (a) plots the estimated interpretation wedge $-\widehat{\mu}_c^k$, where higher values correspond to a more favorable mapping from the common topic signal $\widehat{L}_t^k$ into sentiment. Panel (b) plots the estimated updating intensity $\widehat{\kappa}_c^k \equiv 1 - \widehat{\lambda}_c^k$, the average weight on new public information. The public signal $\widehat{L}_t^k$ is the topic-k common factor (first PC across countries) normalized within topic.

(a) LS-style ΔR^2 Method



(b) Shapley Decomposition Method

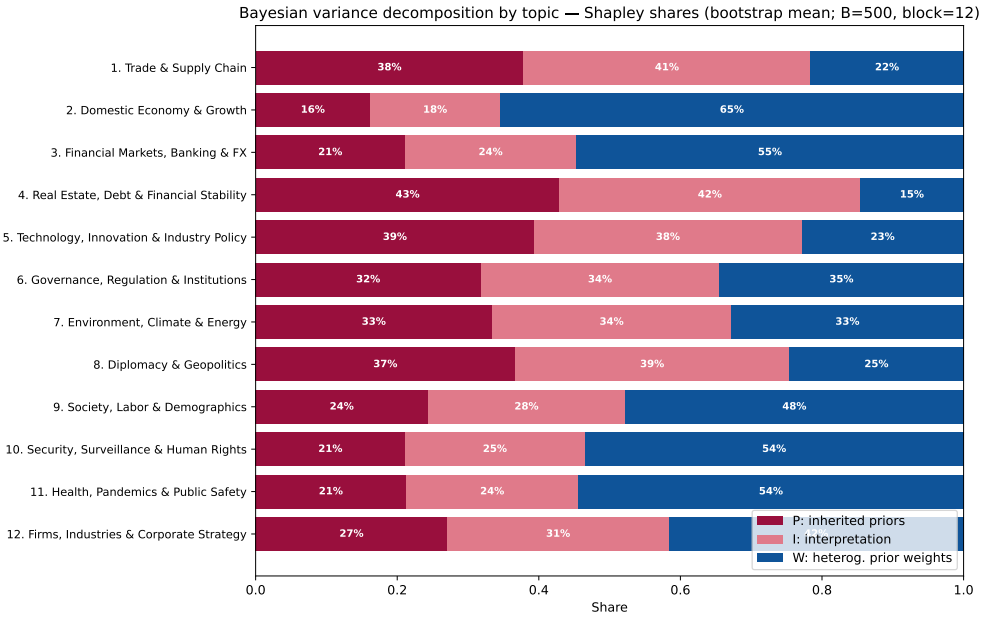


Figure 8: Belief Level vs. Updating Heterogeneity: Variance Decomposition of Topic Sentiment Index

Notes: This figure presents a static variance decomposition of cross-country disagreement in media sentiment toward China across twelve topic categories. The decomposition attributes cross-country dispersion to two broad mechanisms. First, *belief-level disagreement*, which includes (i) inherited baseline differences in stance (priors, P) and (ii) heterogeneous interpretation of common public information (I). Second, *updating disagreement*, reflecting cross-country differences in the weight placed on prior beliefs relative to new information (updating intensity, W). The top panel reports results using the LS-style ΔR^2 approach, which measures the marginal contribution of each component to the time-series variation in observed cross-sectional variance. The bottom panel reports results using an order-invariant Shapley decomposition, which allocates the model-implied variance surplus across components by averaging marginal contributions over all possible inclusion orders.

Table 1: Summary statistics

Panel A: Narrative indices								
	mean	**sd**	**p5**	**p25**	**p50**	**p75**	**p95**	**count**
(1) Aggregate-level Index:								
$num_{c,t}$	0.10	0.07	0.02	0.04	0.07	0.14	0.24	2587
$sen_{c,t}$	-0.09	0.12	-0.27	-0.17	-0.11	-0.02	0.14	2587
$sen^{bag}_{c,t}$ (×100)	-0.96	0.42	-1.63	-1.23	-0.96	-0.70	-0.23	2587
$risk_{c,t}$ (×100)	0.16	0.04	0.09	0.13	0.15	0.18	0.23	2587
$pos_{c,t}$	0.01	0.00	0.01	0.01	0.01	0.01	0.01	2587
$neg_{c,t}$	0.02	0.00	0.01	0.02	0.02	0.02	0.02	2587
(2) Topic-level Index: Topic Share Index ($share^k_{c,t}$):								
1.Trade & Supply Chain	0.09	0.04	0.04	0.06	0.08	0.11	0.17	2587
2.Domestic Economy & Growth	0.08	0.04	0.03	0.06	0.08	0.10	0.14	2587
3.Financial Markets, Banking & FX	0.10	0.08	0.02	0.04	0.08	0.13	0.27	2587
4.Real Estate, Debt & Financial Stability	0.02	0.02	0.00	0.01	0.02	0.03	0.05	2587
5.Technology, Innovation & Industry Policy	0.03	0.02	0.01	0.02	0.03	0.04	0.06	2587
6.Governance, Regulation & Institutions	0.05	0.03	0.02	0.04	0.05	0.06	0.11	2587
7.Environment, Climate & Energy	0.04	0.03	0.01	0.02	0.04	0.06	0.09	2587
8.Diplomacy & Geopolitics	0.16	0.09	0.06	0.10	0.14	0.20	0.34	2587
9.Society, Labor & Demographics	0.20	0.08	0.07	0.14	0.20	0.25	0.33	2587
10.Security, Surveillance & Human Rights	0.05	0.03	0.01	0.03	0.05	0.07	0.10	2587
11.Health, Pandemics & Public Safety	0.03	0.05	0.00	0.00	0.01	0.02	0.13	2587
12.Firms, Industries & Corporate Strategy	0.15	0.07	0.04	0.10	0.15	0.19	0.27	2587
(3) Topic-level Index: Topic sentiment Index ($sen^k_{c,t}$):								
1.Trade & Supply Chain	-0.03	0.25	-0.41	-0.20	-0.03	0.12	0.37	2609
2.Domestic Economy & Growth	-0.07	0.25	-0.44	-0.23	-0.09	0.06	0.37	2609
3.Financial Markets, Banking & FX	-0.09	0.27	-0.50	-0.25	-0.10	0.04	0.38	2609
4.Real Estate, Debt & Financial Stability	-0.08	0.31	-0.56	-0.24	-0.05	0.04	0.46	2609
5.Technology, Innovation & Industry Policy	0.08	0.22	-0.23	-0.03	0.03	0.19	0.47	2609
6.Governance, Regulation & Institutions	-0.20	0.18	-0.48	-0.31	-0.21	-0.09	0.08	2609
7.Environment, Climate & Energy	-0.14	0.23	-0.48	-0.27	-0.14	0.00	0.23	2609
8.Diplomacy & Geopolitics	-0.13	0.17	-0.38	-0.25	-0.14	-0.01	0.15	2609
9.Society, Labor & Demographics	-0.11	0.11	-0.29	-0.17	-0.11	-0.05	0.08	2609
10.Security, Surveillance & Human Rights	-0.37	0.20	-0.67	-0.49	-0.39	-0.26	0.00	2609
11.Health, Pandemics & Public Safety	-0.18	0.29	-0.74	-0.33	-0.13	0.00	0.22	2609
12.Firms, Industries & Corporate Strategy	0.08	0.19	-0.18	-0.03	0.06	0.19	0.41	2609

Panel B: Other Variables								
	mean	**sd**	**p5**	**p25**	**p50**	**p75**	**p95**	**count**
Ret_t	0.00	0.08	-0.13	-0.04	0.00	0.04	0.13	185
Vol_t	0.02	0.01	0.01	0.01	0.01	0.02	0.03	185
$Flow_{ic,t}$	7.36	49.19	-32.56	-4.67	0.00	6.92	58.62	344253
$Growthdiff_{c,t}$	5.08	4.66	-1.23	2.83	4.75	7.30	12.52	915
$EX_{c,t}$	3.34	3.37	0.01	0.21	1.85	5.90	9.45	915
$Intdiff_{c,t}$	0.72	2.88	-5.09	-0.80	0.97	2.74	5.13	903
$Retdiff_{c,t}$	0.00	0.04	-0.07	-0.03	0.00	0.02	0.10	915
$Volratio_{c,t}$	1.68	0.76	0.76	1.12	1.53	2.03	3.17	915

Notes: This table reports summary statistics for all variables used in the empirical analysis. All media indices are measured at a monthly frequency and cover the period from January 2007 to May 2022. For each country c and month t, $num_{c,t}$ denotes the news attention index on China, and $sen_{c,t}$ denotes the aggregate media sentiment index on China. The variable $sen^{bag}_{c,t}$ is a sentiment index constructed using the bag of words method, while $risk_{c,t}$ captures country level risk perceptions about China reflected in domestic media coverage. The variables $pos_{c,t}$ and $neg_{c,t}$ measure positive and negative media narratives, respectively. The topic share index $share^k_{c,t}$ characterizes the share of topic k among all China related news covered by media outlets in country c in month t, and $sen^k_{c,t}$ measures the sentiment associated with topic k. Stock return, denoted by Ret_t, is measured as the monthly average of daily returns on the Shanghai Shenzhen CSI 300 index. Stock market volatility, denoted by Vol_t, is computed as the monthly volatility of daily returns on the same index. The variable $Flow_{ic,t}$ represents quarterly investment flows of fund i domiciled in country c into Chinese assets, as described in Section 6. The variable $Growthdiff_{c,t}$ is the year over year gross domestic product growth differential between China and the investor's domicile economy. The bilateral exchange rate between the Chinese renminbi and the investor's currency is denoted by $EX_{c,t}$. The variable $Intdiff_{c,t}$ measures the interest rate differential between China and the investor's domicile economy. The variable $Retdiff_{c,t}$ captures the stock market return differential between China and the investor's domicile economy, while $Volratio_{c,t}$ denotes the ratio of stock market volatility in China relative to that of the investor's domicile economy. Detailed definitions of all macroeconomic and financial variables are provided in Appendix E.

Table 2: The origins of national belief levels

Dep. Var.:	$-\widehat{\mu}_c^k$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Economic interdependence with China:								
$TradeCN_c$	0.398							0.399*
	(0.192)							-0.143
$TradeShare_c$	-0.128							0.468**
	(0.135)							-0.111
Geopolitical alignment:								
$AgreeCN_c$		0.207**						
		(0.064)						
BRI_c		-0.017						-0.115***
		(0.024)						(0.013)
Proximity:								
$DistCN_c$			-0.047					0.000
			(0.026)					(0.005)
$ContCN_c$			0.030					-0.123***
			(0.026)					(0.023)
$ComLangCN_c$			0.033					
			(0.055)					
Institutional quality:								
$RuleLaw_c$				-0.049*				-0.076***
				(0.020)				(0.007)
$GovEff_c$					-0.045			
					(0.023)			
$RegQuality_c$						-0.047*		
						(0.018)		
Market depth and financial similarity:								
$ListedFirms_c$							-0.000	-0.001***
							(0.000)	0.000
$Turnover_c$							-0.000	-0.001*
							(0.000)	0.000
$RetCorrCN_c$							0.334	0.717***
							(0.248)	(0.099)
Constant	-0.104**	-0.183**	0.332	-0.016	-0.016	-0.015	-0.106	-0.089***
	(0.024)	(0.043)	(0.232)	(0.028)	(0.035)	(0.029)	(0.088)	(0.001)
Topic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.405	0.421	0.424	0.427	0.379	0.407	0.358	0.538
N	168	168	168	168	168	168	156	156

Notes: This table reports estimates of equation (10). The dependent variable is the estimated country-specific narrative belief level across 12 topics. The independent variables capture five conceptual dimensions: economic interdependence with China, geopolitical alignment, geographic proximity, institutional quality, and market depth and financial similarity. All country characteristics are measured as averages over the period 2007–2022, unless otherwise noted due to data availability. Many country-level characteristics, particularly those capturing institutional quality and geopolitical alignment, are highly correlated. Including all alternative measures simultaneously would induce multicollinearity in a small cross-country sample. We therefore include one or two representative measures per conceptual category in column (8). Appendix Table A7 reports the full correlation matrix. Standard errors are two-way clustered at the country and topic levels and are reported in parentheses. Statistical significance is indicated by *, **, and *** for the 10%, 5%, and 1% significance levels, respectively.

Table 3: Baseline: media narratives and flows

Dep. Var.:	$Flow_{ic,t}$			
	(1)	(2)	(3)	(4)
$sen_{c,t-1}$	1.042***	1.047***	0.977***	0.955***
	(0.245)	(0.245)	(0.247)	(0.247)
$num_{c,t-1}$	0.210	0.141	0.270	0.216
	(0.258)	(0.259)	(0.257)	(0.258)
$Growthdiff_{c,t-1}$		0.039		0.026
		(0.060)		(0.061)
$EX_{c,t-1}$		-1.535***		-1.735***
		(0.321)		(0.328)
$Intdiff_{c,t-1}$		-0.339		-0.414
		(0.222)		(0.227)
$Retdiff_{c,t-1}$			22.560*	26.051**
			(9.857)	(9.924)
$Volratio_{c,t-1}$			0.862*	1.345**
			(0.407)	(0.419)
$Constant$	7.405***	19.087***	6.159***	18.845***
	(0.064)	(2.676)	(0.601)	(2.671)
Fund FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Adj. R^2	0.055	0.055	0.055	0.055
N	307603	307541	307603	307541

Notes: This table reports results from equation (11) estimated at the fund-quarter level over the period 2007Q1-2022Q1 for a sample of funds domiciled in 16 countries. The dependent variable is the percentage change in the flow of institutional investors' investments in China. The main independent variable is the sentiment index, which is standardized for each country. Columns (2) and (4) include a set of macroeconomic controls that may influence cross-border flows: the year-on-year real GDP growth differential between China and the investor's domicile economy, the bilateral exchange rate between the Chinese renminbi and the currency of the investor's domicile country, and the interest rate differential between China and the investor's economy. Columns (3) and (4) include a set of financial controls: the return differential between the Chinese stock market and that of the investor's domicile economy, and the volatility ratio of the Chinese stock market and that of the investor's domicile economy, measured by the volatility of the Shanghai Shenzhen CSI 300 index. All independent variables are lagged by one quarter. All columns include fund fixed effects and quarter fixed effects. Standard errors are clustered at the fund level and are reported in parentheses. Statistical significance is indicated by *, **, and *** for the 10%, 5%, and 1% significance levels, respectively.

Table 4: Robustness checks: media narratives and flows

Panel A: Bag-of-word media narratives and flows				
Dep. Var.:	$Flow_{ic,t}$			
	(1)	(2)	(3)	(4)
$sen^{bag}_{c,t-1}$	0.945***	0.963***	0.913***	0.933***
	(0.260)	(0.260)	(0.261)	(0.261)
$num_{c,t-1}$	0.122	0.065	0.206	0.170
	(0.250)	(0.252)	(0.249)	(0.252)
Macro Controls	No	Yes	No	Yes
Financial Controls	No	No	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Adj. R^2	0.055	0.055	0.055	0.055
N	307603	307541	307603	307541
Panel B: Cross-language media narratives and flows				
Dep. Var.:	$Flow_{ic,t}$			
	(5)	(6)	(7)	(8)
$csen_{c,t-1}$	0.027	0.817**	0.803**	0.746**
	(0.216)	(0.250)	(0.252)	(0.251)
$cnum_{c,t-1}$	-0.016	0.141	0.253	0.199
	(0.231)	(0.264)	(0.262)	(0.263)
Macro Controls	No	Yes	No	Yes
Financial Controls	No	No	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Adj. R^2	0.055	0.057	0.057	0.057
N	314009	278263	278325	278263
Panel C: Media narrative shocks and flows				
Dep. Var.:	$Flow_{ic,t}$			
	(9)	(10)	(11)	(12)
$senshock_{c,t-1}$	1.574***	1.474***	1.579***	1.470***
	(0.311)	(0.314)	(0.313)	(0.315)
$num_{c,t-1}$	0.008	-0.050	0.092	0.050
	(0.248)	(0.250)	(0.247)	(0.249)
Macro Controls	No	Yes	No	Yes
Financial Controls	No	No	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Adj. R^2	0.054	0.054	0.054	0.054
N	306347	306285	306347	306285

Notes: This table reports estimates of equation (11) using fund quarter level data over the period from 2007Q1 to 2022Q1 for a sample of institutional funds domiciled in 16 countries. The dependent variable is the percentage change in institutional investors' investment flows into Chinese assets. The main independent variable is the media narrative sentiment measure, constructed using the bag of words method in Panel A as described in Section 6.1, the cross language media sentiment index in Panel B as described in Section 6.2, and media narrative shocks in Panel C as described in Section 6.3. Columns (2), (4), (6), (8), (10), and (12) include a set of macroeconomic controls that may influence cross border investment flows. These controls include the year on year real gross domestic product growth differential between China and the investor's domicile economy, the bilateral exchange rate between the Chinese renminbi and the currency of the investor's domicile country, and the interest rate differential between China and the investor's economy. Columns (3), (4), (7), (8), (11), and (12) additionally include financial controls, namely the stock market return differential between China and the investor's domicile economy and the ratio of stock market volatility in China relative to that of the investor's domicile economy, measured using the volatility of the Shanghai Shenzhen CSI 300 index. All independent variables are lagged by one quarter. All specifications include fund fixed effects and quarter fixed effects. Standard errors are clustered at the fund level and reported in parentheses. Statistical significance is indicated by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Table 5: Media narratives and capital flows following the Arab Spring

Panel A: Stage 1 : Media Narrative Changes following Arab Spring					
(a) Attention Index:					
Dep. Var.:			$num_{c,t}$		
Event Window:	12m	18m	24m	30m	36m
	(1)	(2)	(3)	(4)	(5)
$Post \times HighGeoPolitics$	-0.003	0.003	0.009	0.014	0.012
	(0.006)	(0.005)	(0.007)	(0.011)	(0.010)
Country FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.982	0.978	0.964	0.949	0.943
N	336	504	668	820	971
(b) Sentiment index:					
Dep. Var.:			$sen_{c,t}$		
Event Window:	12m	18m	24m	30m	36m
	(6)	(7)	(8)	(9)	(10)
$Post \times HighGeoPolitics$	-0.037*	-0.047**	-0.040**	-0.034***	-0.034**
	(0.016)	(0.015)	(0.010)	(0.007)	(0.008)
Country FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.757	0.770	0.774	0.764	0.752
N	336	504	668	820	971
Panel B: Stage 2: Fund Flow Changes following Arab Spring					
Dep. Var.:			$Flow_{ic,t}$		
Event Window:	4q	6q	8q	10q	12q
	(11)	(12)	(13)	(14)	(15)
$Post \times HighGeoPolitics$	-6.103***	-5.090***	-4.782***	-3.851***	-4.030***
	(1.589)	(1.330)	(1.209)	(1.154)	(1.113)
Fund FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Macro & Financial Controls	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.074	0.069	0.067	0.069	0.064
N	32434	46084	58694	71042	83492

Notes: This table reports results from an event study analysis that uses the Arab Spring as a global narrative shock, as described in Section 6.4. Panel A examines changes in media attention to China in panel (a) and media sentiment toward China in panel (b) for countries with high geopolitical attention following the Arab Spring event, relative to countries with low geopolitical attention, as specified in equation (12). Panel B examines changes in institutional investment flows for countries with high media sentiment exposure relative to countries with low exposure, as specified in equation (13). Each column corresponds to a different event window before and after the Arab Spring event. Standard errors are clustered at the country level in Panel A and at the fund level in Panel B, and are reported in parentheses. Statistical significance is indicated by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Table 6: Media narratives and capital flows using the shift-share method

Panel A: Second Stage (Flows)				
Dep. Var.:	$Flow_{ic,t}$			
	(1)	(2)	(3)	(4)
$\widehat{sen}^{IV}_{c,t-1}$	1.537***	1.472**	1.463**	1.367**
	(0.467)	(0.467)	(0.468)	(0.469)
$num_{c,t-1}$	0.018	-0.049	0.102	0.052
	(0.250)	(0.252)	(0.249)	(0.251)
Macro Controls	No	Yes	No	Yes
Financial Controls	No	No	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Adj. R^2	0.055	0.055	0.055	0.055
N	307603	307541	307603	307541
Panel B: First Stage (Monthly)				
$Z_{c,t}$	0.900**			
	(0.223)			
Country FE	Yes			
Month FE	Yes			
First-stage F-stat	16.26			
Adj. R^2	0.686			
N	2486			

Notes: This table reports estimates from the shift-share identification strategy used to construct an exogenous media narrative shock and assess its impact on capital flows, as described in Section 7. Panel A presents the second-stage estimates of the effect of instrumented media sentiment on portfolio flows, corresponding to equation(18), under alternative specifications that sequentially add macroeconomic and financial control variables. Panel B reports the corresponding first-stage results from equation(17), which relate domestic media sentiment to the shift–share instrument constructed from non-China WSJ narrative shocks using the monthly data. Standard errors are clustered at the fund level in Panel A and at the investor-country level in Panel B and are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Table 7: Asymmetric and higher-moment effects of media narratives on capital flows

Panel A: Asymmetric effects of media sentiment on flows				
Dep. Var.:	$Flow_{ic,t}$			
	(1)	(2)	(3)	(4)
$pos_{c,t-1}$	0.290	0.346	0.345	0.422
	(0.211)	(0.219)	(0.212)	(0.219)
$neg_{c,t-1}$	-0.775**	-0.770**	-0.712**	-0.694**
	(0.252)	(0.255)	(0.253)	(0.256)
$num_{c,t-1}$	0.133	0.082	0.223	0.195
	(0.250)	(0.253)	(0.250)	(0.253)
Macro Controls	No	Yes	No	Yes
Financial Controls	No	No	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Adj. R^2	0.055	0.055	0.055	0.055
N	307603	307541	307603	307541
Panel B: Second-moment (risk) effects of media narratives on flows				
Dep. Var.:	$Flow_{ic,t}$			
	(5)	(6)	(7)	(8)
$risk_{c,t-1}$	-0.817**	-0.667**	-0.825**	-0.659**
	(0.251)	(0.250)	(0.251)	(0.250)
$num_{c,t-1}$	0.194	0.084	0.290	0.192
	(0.262)	(0.261)	(0.261)	(0.261)
Macro Controls	No	Yes	No	Yes
Financial Controls	No	No	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Adj. R^2	0.055	0.055	0.055	0.055
N	307603	307541	307603	307541

Notes: This table reports estimates of equation (11) using fund quarter level data over the period from 2007Q1 to 2022Q1 for a sample of institutional funds domiciled in 16 countries. The main independent variables are the positive and negative media sentiment indices in Panel and the media risk index in Panel B, as described in Section 8. The positive index, negative index, and risk index are standardized within each country. Columns (2), (4), (6), and (8) include a set of macroeconomic controls that may affect cross border investment flows. These controls include the year on year real gross domestic product growth differential between China and the investor's domicile economy, the bilateral exchange rate between the Chinese renminbi and the currency of the investor's domicile country, and the interest rate differential between China and the investor's economy. Columns (3), (4), (7), and (8) additionally include financial controls, namely the stock market return differential between China and the investor's domicile economy and the ratio of stock market volatility in China relative to that of the investor's domicile economy, measured using the volatility of the Shanghai Shenzhen CSI 300 index. All independent variables are lagged by one quarter. All specifications include fund fixed effects and quarter fixed effects. Standard errors are clustered at the fund level and reported in parentheses. Statistical significance is indicated by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Appendix for Online Publication

Table of Contents

A Media List **1**

B Additional Indices **2**

C Sentiment Construction: Examples **4**

D Cross-Language FinBERT Model Selection **5**
D.1 English 5
D.2 German 5
D.3 French 6
D.4 Spanish 6

E Variable Definitions **7**

F Prompt Used for Subject Classification **11**

G Bayesian Learning Method Derivation **12**
G.1 Deriving the disagreement recursion 12
G.2 Estimation: Variance decomposition 13
G.2.1 Estimation of λ via the forecast-revision regression 14
G.2.2 Estimation of σ_λ^2 and σ_μ^2 15
G.3 Counterfactual variance paths and channel attribution 16

H Correlations of National Belief Level Variables **20**

A Media List

Table A1: Media List

Region abbr	Region	Source	Date	Additional excluded months
AU	Australia	The Australian	Jan 01, 2007 to May 31, 2022	
AU	Australia	The Australian Financial Review	Sep 02, 2013 to May 31, 2022	Jun 2015 (*)
AU	Australia	The Canberra Times	Jan 01, 2007 to May 31, 2022, with a gap from Jan 1, 2010-Oct 21, 2010	
AU	Australia	Sydney Morning Herald	Jan 01, 2007 to May 31, 2022	
CA	Canada	The Globe and Mail	Jan 01, 2007 to May 31, 2022	
CA	Canada	Montreal Gazette	Jan 01, 2007 to May 31, 2022	
CA	Canada	The Vancouver Sun	Jan 01, 2007 to May 31, 2022	
CA	Canada	National Post	Jan 01, 2007 to May 31, 2022	
CA	Canada	Toronto Star	Jan 01, 2007 to May 31, 2022	
EU	EMU	Financial Times	Jan 01, 2007 to May 31, 2022	
HK	Hong Kong	The Wall Street Journal Asia	Jan 02, 2007 to Oct 06, 2017	
HK	Hong Kong	China Daily (Hong Kong ed.)	Jul 22, 2013 to May 31, 2022, with a gap from Aug 21, 2016-Aug 23,2017, a gap from Jan 19, 2019-April 30, 2019, and a gap from Jan 2, 2020-Jan 2, 2022	May 2015 (*)
IN	India	The Times of India	Jan 01, 2007 to May 31, 2022	
IN	India	The Hindustan Times	Jan 01, 2007 to May 31, 2022	
IN	India	Indian Express	Apr 23, 2009 to May 31, 2022	
IE	Ireland	Irish Times	Jan 02, 2007 to May 31, 2022	
IE	Ireland	Sunday Independent	Jan 07, 2007 to May 29, 2022,	
IE	Ireland	Irish Independent	Jan 04, 2007 to May 30, 2022,	
MY	Malaysia	New Straits Times	Jan 01, 2007 to May 31, 2022	
NZ	New Zealand	The New Zealand Herald	Jan 01, 2007 to May 07, 2022, with a gap from Dec 13, 2013-Jan 13,2015	
PH	Philippine	Business Mirror	Jan 01, 2014 to May 31, 2022	
SG	Singapore	The Straits Times	Jan 01, 2011 to May 31, 2022	
SG	Singapore	The Business Times	Jan 01, 2011 to May 31, 2022	
ZA	South Africa	The Mercury	March 31, 2008 to May 31, 2022 with a gap from Oct 13, 2012-May 1, 2015	Aug 2008, Sep 2008, Oct 2008, Nov 2008
ZA	South Africa	The Star	Mar 31, 2008 to May 31, 2022, with a gap from Oct 14, 2012-Dec 8, 2014	Aug 2008, Sep 2008, Oct 2008, Nov 2008
KR	South Korea	The Korea Times	Apr 04, 2007 to May 31, 2022	
TW	Taiwan	China Post	Sep 06, 2011 to Oct 03, 2017	Aug 2017, Sep 2017
TH	Thailand	Asia News Monitor	Jul 30, 2008 to May 31, 2022, with a gap from Aug 8, 2008-Dec 31, 2008 with a gap from Dec 30, 2010-July 19, 2011	July 2008 (*)
TH	Thailand	The Nation	Jan 23, 2012 to Apr 29, 2020 with a gap from Mar 12, 2019-Sep 30, 2019	Dec 2017,
UK	U.K.	Daily Mail	Jan 01, 2007 to May 31, 2022	
UK	U.K.	The Daily Telegraph	Jan 01, 2007 to May 31, 2022	
UK	U.K.	The Daily Mirror	Jan 01, 2007 to May 31, 2022	
UK	U.K.	Evening Standard	Jan 02, 2007 to May 31, 2022	
UK	U.K.	Financial Times	Jan 01, 2007 to May 31, 2022	
US	U.S.	Wall Street Journal	Jan 01, 2007 to May 31, 2022	
US	U.S.	New York Times	Jan 01, 2007 to May 31, 2022	
US	U.S.	The Washington Post	Jan 01, 2007 to May 31, 2022	
US	U.S.	USA Today	Jan 01, 2007 to May 31, 2022	
US	U.S.	Boston Globe	Jan 01, 2007 to May 31, 2022	
US	U.S.	The Los Angeles Times	Jan 01, 2007 to May 31, 2022	

Notes: This table shows the list of newspapers in each country used to construct all media indices. If a month is marked with * in the additional excluded months column, it means that month is not excluded, despite the total number of news articles being relatively lower than in adjacent months.

Table A2: Media List (Non-English Newspapers)

Region abbr	Region	Source	Date	Language	Keyword
DE	Germany	Die Tageszeitung	Aug 5, 2008-May 31, 2022	German	China
DE	Germany	Die Welt	Nov 1, 2009-May 31, 2022	German	China
DE	Germany	Welt am Sonntag	Jan 6, 2008-May 31, 2022	German	China
FR	France	Le Monde	Jan 1, 2007-May 31, 2022	French	Chine
ES	Spain	El Pais	July 28, 2008-May 31, 2022	Spanish	China
CH	Switzerland	Le Temps	Sep 22, 2010-May 31, 2022	French	Chine
AT	Austria	Die Presse	May 28,2011-May 31, 2022	German	China

Notes: This table shows the list of newspapers in each country used to construct all media indices in language other than English.

B Additional Indices

For each media outlet, we conduct a textual analysis to compile the raw data used for the following indices, including: (1) The total number of China-related news articles each month t in each newspaper m ($num_{m,t}$); (2) The total number of news articles in each newspaper each month ($all_{m,t}$); (3) The total number of words in each China-related article i for each newspaper m on date d ($totalwords_{im,d}$); (4) The total negative word count in each article ($negsum_{im,d}$), using the negative word list developed by Loughran and McDonald (2011) from firm 10-K filings, which contains 2,337 words; (5) The total positive word count in each article ($possum_{im,d}$), using the positive word list in Loughran and McDonald (2011), which includes 353 words.(6) The total count of risk-related words in each article ($risksum_{im,d}$) using the risk word list in Hassan, Hollander, Van Lent and Tahoun (2019), which includes all single-word synonyms of "risk", "risky", and "uncertainty" as listed in the Oxford Dictionary (excluding "question", "questions", and "venture"). This risk word list comprises 123 words.

Bag-of-word Method Sentiment Index. We follow the similar methodology of Flynn and Sastry (2024) and Hassan, Hollander, Van Lent and Tahoun (2019) to construct the sentiment index ($sen_{c,t}^{bag}$) for investors domiciled in country c in month t, which serves

as one of our key measures of media narratives at the intensive margin. For each article, we first calculate sentiment as the difference between the number of positive words ($possum_{im,d}$) and the number of negative words ($negsum_{im,d}$), scaled by the total number of words in the article ($totalwords_{im,d}$) to control for differences in article length. We then compute the average article sentiment for each media outlet in each month. Finally, the country-level sentiment index is obtained by averaging these media-level sentiment measures across all outlets within each country.

Risk Index. We follow the approach in Hassan, Hollander, Van Lent and Tahoun (2019) to construct the risk index ($risk_{c,t}$) for investors domiciled in country c at month t. For each article, we compute the share of risk-related words by dividing the total number of risk words ($risksum_{im,d}$) by the total word count of the article ($totalwords_{im,d}$), which adjusts for differences in article length. We then take the monthly average of these risk shares within each media outlet. The country-level risk index is constructed by aggregating these outlet-level averages across all newspapers within a country.

C Sentiment Construction: Examples

Table A3: Examples of Sentiment Classification

Sentiment	Conf.	Source	Date	Title and Excerpt
Negative	0.9355	US, *The Washington Post*	May 24, 2015	*Beijing: U.S. flight 'very irresponsible'.* "China accused the United States on Friday of "very irresponsible" and potentially dangerous actions after a U.S. surveillance plane flew near disputed islands in the South China Sea and was warned off by the Chinese navy this week....Experts said China appears to be trying to create a military exclusion zone around the islands, a move that could heighten regional tensions."
Negative	0.9088	US, *Wall Street Journal*	Feb 12, 2012	*Corporate News: Lead-Pollution Fight in China — Shanghai Authorities Won't Allow Johnson Controls to Resume Battery Output.* "Authorities here directly linked lead pollution that they said had sickened local children to emissions from a Johnson Controls Inc. battery plant and indicated the facility would not be permitted to process lead in the future. The move comes amid rising worries in China over pollution...It also raises issues of land use – a persistent problem in the fast-growing country... "
Neutral	0.8996	SG, *The Straits Times*	Jun 28, 2018	*Beijing spent $65b on charm offensive to win friends: US study.* "China has gone all out to woo its neighbours in recent years, spending at least US$48 billion (S$65 billion) to win friends between 2000 and 2016, a US study has found. The lion's share, or US$45.8 billion, was spent on infrastructure, while the rest went to humanitarian aid, debt relief, and grants and loans given as "gifts" with no earmarked purpose..."
Neutral	0.6514	US, *The Washington Post*	May 31, 2022	*How the U.N. handed China a propaganda victory.* "Before U.N. High Commissioner for Human Rights Michelle Bachelet made her long-awaited tip to China last week, the Biden administration and the human rights community urged her not to let Beijing turn the visit into a propaganda win for the Chinese Communist Party. But Bachelet ignored those warnings. Her trip ended up helping China deny its genocide against Uyghur Muslims and other repressive policies, harming the cause of human rights accountability in the process..."
Positive	0.8718	US, *Wall Street Journal*	Nov 14, 2017	*China Leads Charge As IPOs Roar Back.* "A flood of Chinese companies is driving the biggest world-wide surge of initial public offerings in a decade. More than 1,450 companies globally have gone public so far in 2017, putting this year on track to become the busiest for new listings since 2007, according to Dealogic data through Friday. Roughly two-thirds of the IPOs were in the Asia-Pacific region, which has roared past the U.S. to become the dominant region for new stock listings..."
Positive	0.8711	CA, *National Post*	Mar 25, 2015	*Emerging markets cause a rethink; Resurgence offers investors top returns.* "Emerging market stocks are dogs. At least that's what a lot of investors have come to believe in recent years, but their total shareholder returns over the past year suggests just the opposite is true. Led by a resurgence in Chinese stocks, eight of the top 10 equity indexes in the world were located in developing markets in 2014, and Asia's top benchmark, MSCI Asia-Pacific Index, has continued its recent run this year. It rose for the fifth straight day on Tuesday, and it is now up 8.26% so far in 2015..."

Notes: This table presents representative examples of article-level sentiment classifications produced by the language model. The reported confidence score corresponds to the model's posterior probability assigned to the predicted sentiment label. Article titles are shown in italics, followed by a representative excerpt used for classification.

D Cross-Language FinBERT Model Selection

This appendix summarizes the model-selection criteria used in our multilingual FinBERT sentiment analysis pipeline. For each language, we review available FinBERT-style transformer models and document the final package adopted in our analysis. We focus on conceptual strengths and limitations of each model without presenting sentence-level comparisons or numerical evaluation results. All selected packages follow the standard three-way polarity structure (`positive`, `neutral`, `negative`), ensuring consistency across languages.

D.1 English

Selected Package: `ProsusAI/finbert`

Among multiple English-language FinBERT variants, we adopt `ProsusAI/finbert` as the primary sentiment engine in our pipeline. The model is fine-tuned on a large corpus of financial documents, allowing it to capture tone in macroeconomic policy discussions, corporate disclosures, and market commentary. Compared with alternative packages, it produces better-calibrated confidence distributions and avoids overly sharp classifications.

A further advantage is its compatibility with multilingual sentiment workflows: its labeling structure aligns closely with FinBERT-style models in other languages. A potential limitation is that its sensitivity to environmental or policy-oriented sentiment can be weaker than that of some non-English domain-specific models, reflecting the narrower scope of its fine-tuning data.

D.2 German

Selected Package: `Scherrmann/GermanFinBERT`

For German financial sentiment classification, we select `Scherrmann/GermanFinBERT`. This model is explicitly trained on German financial disclosures, earnings reports, and

economic news, enabling it to accurately detect tone in both firm-level and macroeconomic statements. It distinguishes nuanced sentiment shifts—particularly in ambiguous or mixed-tone contexts—more effectively than general-purpose German sentiment models.

The model's native German labels (`Positiv`, `Neutral`, `Negativ`) map cleanly into the FinBERT framework used across languages. Its primary limitation is conservative treatment of non-financial domains such as political or environmental reporting, for which it sometimes assigns more neutral sentiment than expected.

D.3 French

Selected Package: `bardsai/finance-sentiment-fr-base`

For French-language analysis, we adopt the financial-domain model `bardsai/finance-sentiment-fr-base`. This DistilCamemBERT-based transformer is fine-tuned on labeled French financial sentences, allowing it to reliably classify tone in earnings releases, policy announcements, and macroeconomic commentary. It also generalizes well to adjacent domains relevant for financial research, such as regulatory developments and environmental reporting.

The model retains the standard FinBERT polarity scheme, ensuring comparability with other languages in our pipeline. As a distilled model, it may capture fewer subtle linguistic nuances than a full CamemBERT-based architecture, but its efficiency and stability make it well-suited for large-scale text processing.

D.4 Spanish

Selected Package: `bardsai/finance-sentiment-es-base`

For Spanish financial sentiment classification, we select `bardsai/finance-sentiment-es-base`. Unlike general-purpose Spanish sentiment models, this package is fine-tuned specifically on financial text, yielding greater sensitivity to credit risk language, earnings-related

tone, fiscal and monetary policy narratives, and broader macro-financial developments.

The model also performs consistently in political or environmental contexts when such narratives influence financial markets. Its FinBERT-compatible labeling scheme facilitates seamless integration into the multilingual pipeline. A potential limitation is that it may underperform general-purpose models in highly colloquial or conversational text, although such cases are not central to financial analysis.

In summary, the selected models share key advantages: strong financial-domain alignment, well-calibrated sentiment distinctions, and consistent polarity structures across languages. Together, they provide a unified multilingual FinBERT framework for cross-country sentiment analysis in financial research.

E Variable Definitions

Table A4 shows the definitions and source of variables used in equation (11).

Table A4: Variable definitions

Variable	Def	Source
$Growthdiff_{c,t}$	GDP growth difference between China and invstero domicle country, where GDP is the real quarterly GDP growth on a year-to-year basis	CEIC
$EX_{c,t}$	The exchange rate between the Chinese renminbi and the investor's domicile currency, expressed as the number of renminbi required to purchase one unit of the domicile country's currency.	Bloomberg
$Intdiff_{c,t}$	Interest rate differential between China and the fund domicile country, which is constructed using short-term interest rates from the *OECD*. In cases where a country's short-term interest rates are unavailable, we substitute with the lending interest rate difference from the *IMF*. If both sets of rates are missing, we default to using China's short-term interest rates.	OECD, IMF International Financial Statistics (IFS)
$Retdiff_{c,t}$	The stock market return differential is defined as the difference between China's stock market return and that of the investor's domicile country. China's stock market return is measured as the quarterly average of monthly returns on the Shanghai Shenzhen CSI 300 index. The indices used to calculate stock returns for the investor's domicile countries are listed in Table A5.	Bloomberg
$Volratio_{c,t}$	Stock market volatility ratio is defined as the ratio of China's stock market volatility to that of the investor's domicile country. China's stock market volatility is measured as the quarterly average of the monthly volatility of daily returns on the Shanghai Shenzhen CSI 300 index. For the investor's domicile countries, the stock indices used to calculate volatility are listed in Table A5.	Bloomberg

Notes: This table shows the definitions and source of variables used in equation (11).

Table A5: Variable definitions: Bloomberg tickers

Country Code	Index Name	Bloomberg Ticker	Currency Name	Bloomberg Ticker
AU	S&P/ASX 200	AS51	Australian Dollar	AUDCNY
CA	S&P/TSX Composite Index	SPTSX	Canadian Dollar	CADCNY
UK	FTSE All-Share Index	ASX	British Pound	GBPCNY
HK	Hang Seng Index	HSI	Hong Kong Dollar	HKDCNY
Eurozone	Euro Stoxx 50	SX5E	Euro	EURCNY
IN	Nifty 50	NIFTY	Indian Rupee	INRCNY
KR	KOSPI Composite Index	KOSPI	South Korean Won	KRWCNY
MY	FTSE Bursa Malaysia KLCI	FBMKLCI	Malaysian Ringgit	MYRCNY
NZ	S&P/NZX 50 Index	NZSE50FG	New Zealand Dollar	NZDCNY
PH	PSEi (Philippine Stock Exchange Index)	PCOMP	Philippine Peso	PHPCNY
SG	Straits Times Index (STI)	STI	Singapore Dollar	SGDCNY
TH	SET Index	SET	Thai Baht	THBCNY
TW	TAIEX (Taiwan Capitalization Weighted Index)	TWSE	Taiwan Dollar	TWDCNY
US	S&P 500	SPX	US Dollar	USDCNY
ZA	FTSE/JSE All Share Index	JALSH	South African Rand	ZARCNY

Notes: This table lists the Bloomberg tickers for the variables described in Table A4.

Table A6: Country belief level variables

Category	Variable	Definition	Source
Economic interdependence with China	$TradeCN_c$	Total trade with China relative to local GDP	World Bank WITS bilateral trade data, IMF World Economic Outlook (WEO)
	$TradeShare_c$	[Sum of exports and imports between China and country] / [Sum of total exports and imports of China]	World Bank WITS bilateral trade data
Geopolitical alignment	$AgreeCN_c$	UN voting alignment with China	Bailey et al. (2017)
	BRI_c	Belt and Road Initiative participation	Nedopil (2025)
Proximity	$ContCN_c$	Contiguity measure using common border with China.	Conte et al. (2022)
	$DistCN_c$	Log distance between the most populated city of China and the other country, in km, bilateral	Conte et al. (2022)
	$ComLangCN_c$	Common language that equals 1 if the country shares common official or primary languages with China	Conte et al. (2022)
Institutional quality	$RuleLaw_c$	Rule of Law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5.	World Bank World Development Indicator
	$GovEff_c$	Government Effectiveness captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5.	World Bank World Development Indicator
	$RegQuality_c$	Regulatory Quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5.	World Bank World Development Indicator
Market depth and financial similarity	$ListedFirms_c$	Market capitalization of listed domestic companies (% of GDP). Market capitalization (also known as market value) is the share price times the number of shares outstanding (including their several classes) for listed domestic companies. Investment funds, unit trusts, and companies whose only business goal is to hold shares of other listed companies are excluded. Data are end of year values.	World Bank World Development Indicator
	$Turnover_c$	Turnover ratio is the value of domestic shares traded divided by their market capitalization. The value is annualized by multiplying the monthly average by 12.	World Bank World Development Indicator
	$RetCorrCN_c$	Return-based similarity measure to China's stock market.	Bloomberg

Notes: This table lists the variables used in the country belief level analysis in Section 5.

F Prompt Used for Subject Classification

```
You are an expert analyst specializing in international finance, macroeconomics,
China-related news, geopolitics, and global affairs.
Your job is to classify NEWS HEADLINES into exactly ONE topic from the taxonomy below.

Classify based ONLY on the headline.

TOPIC TAXONOMY (12 topics):

1. Trade \& Supply Chain
   Exports, imports, tariffs, sanctions, supply-chain disruptions, logistics,
   commodity trade, sourcing.

2. Domestic Economy \& Growth
   GDP, inflation, consumption, investment, unemployment, industrial output,
   macroeconomic performance, stimulus.

3. Financial Markets, Banking \& FX
   Stocks, bonds, currencies, interest rates, central banks, monetary policy,
   capital flows, banking.

4. Real Estate, Debt \& Financial Stability
   Property markets, mortgages, developers, credit risks, financial fragility.

5. Technology, Innovation \& Industry Policy
   Tech companies, semiconductors, AI, telecom, industrial upgrading, automation,
   state-led industrial policy.

6. Governance, Regulation \& Institutions
   Government policy, regulations, administrative actions, leadership decisions,
   institutional reforms.

7. Environment, Climate \& Energy
   Pollution, climate change, renewable energy, fossil fuels, mining, resources.

8. Diplomacy \& Geopolitics
   Foreign relations, great-power politics, military relations, alliances,
   conflicts, global positioning.

9. Society, Labor \& Demographics
   Population trends, migration, education, workforce, inequality, civil society.

10. Security, Surveillance \& Human Rights
   Policing, national security, censorship, surveillance technology, detentions,
   human-rights disputes.

11. Health, Pandemics \& Public Safety
   Disease outbreaks, vaccines, COVID-19, medical systems, public safety.

12. Firms, Industries \& Corporate Strategy
   Corporate decisions, earnings, M&A, product launches, supply/demand conditions,
   business strategy.

Return ONLY a JSON object in this format:
{"topic": "<topic name>"}
```

G Bayesian Learning Method Derivation

G.1 Deriving the disagreement recursion

Sentiment revision. Define the sentiment revision:

$$\Delta S^k_{c,t} \equiv S^k_{c,t} - S^k_{c,t-1} = (1-\lambda^k_{c,t})\Big(L^k_t - \mu^k_{c,t} - S^k_{c,t-1}\Big). \tag{19}$$

Let $\mathbb{E}_c[\cdot]$ and $\mathrm{Var}_c(\cdot)$ denote cross-sectional moments across countries at a fixed (t,k). Define

$$\bar{S}^k_t \equiv \mathbb{E}_c[S^k_{c,t}], \quad \sigma^2_{S|t,k} \equiv \mathrm{Var}_c(S^k_{c,t}), \quad \lambda^k_t \equiv \mathbb{E}_c[\lambda^k_{c,t}], \quad \sigma^2_{\lambda|t,k} \equiv \mathrm{Var}_c(\lambda^k_{c,t}), \tag{20}$$

and similarly $\mu^k_t \equiv \mathbb{E}_c[\mu^k_{c,t}]$ and $\sigma^2_{\mu|t,k} \equiv \mathrm{Var}_c(\mu^k_{c,t})$. Taking cross-sectional expectations in (19) yields the mean revision:

$$\Delta \bar{S}^k_t \equiv \bar{S}^k_t - \bar{S}^k_{t-1} = (1-\lambda^k_t)\Big(L^k_t - \mu^k_t - \bar{S}^k_{t-1}\Big). \tag{21}$$

Equation (21) is the key step that allows us to eliminate the unobserved term $\left(L^k_t - \mu^k_t - \bar{S}^k_{t-1}\right)$ from the disagreement recursion.

Cross-sectional orthogonality assumption. As in Lahiri and Sheng (2008), we assume that within a given (t,k), the cross-sectional draws $\{S^k_{c,t-1}\}_c$, $\{\lambda^k_{c,t}\}_c$, and $\{\mu^k_{c,t}\}_c$ are mutually independent:

$$\mathrm{Cov}_c(S^k_{c,t-1}, \lambda^k_{c,t}) = \mathrm{Cov}_c(S^k_{c,t-1}, \mu^k_{c,t}) = \mathrm{Cov}_c(\lambda^k_{c,t}, \mu^k_{c,t}) = 0. \tag{22}$$

This assumption delivers a tractable variance recursion and a clean channel interpretation.

Variance algebra From (4),

$$\sigma^2_{S|t,k} = \mathrm{Var}_c\Big(\lambda^k_{c,t} S^k_{c,t-1}\Big) + \mathrm{Var}_c\Big((1-\lambda^k_{c,t})(L^k_t - \mu^k_{c,t})\Big) + 2\,\mathrm{Cov}_c\Big(\lambda^k_{c,t} S^k_{c,t-1},\ (1-\lambda^k_{c,t})(L^k_t - \mu^k_{c,t})\Big). \tag{23}$$

Using the identity $\mathrm{Var}(XY) = \mathrm{Var}(X)\mathrm{Var}(Y) + \mathrm{Var}(X)\mathbb{E}(Y)^2 + \mathrm{Var}(Y)\mathbb{E}(X)^2$ under independence, (22) implies:

$$\mathrm{Var}_c(\lambda^k_{c,t} S^k_{c,t-1}) = \sigma^2_{\lambda|t,k}\sigma^2_{S|t-1,k} + \sigma^2_{\lambda|t,k}(\bar{S}^k_{t-1})^2 + \sigma^2_{S|t-1,k}(\lambda^k_t)^2, \tag{24}$$

$$\mathrm{Var}_c\Big((1-\lambda^k_{c,t})(L^k_t - \mu^k_{c,t})\Big) = \sigma^2_{\lambda|t,k}\sigma^2_{\mu|t,k} + \sigma^2_{\lambda|t,k}(L^k_t - \mu^k_t)^2 + \sigma^2_{\mu|t,k}(1-\lambda^k_t)^2, \tag{25}$$

$$\mathrm{Cov}_c\Big(\lambda^k_{c,t} S^k_{c,t-1},\ (1-\lambda^k_{c,t})(L^k_t - \mu^k_{c,t})\Big) = -(L^k_t - \mu^k_t)\,\bar{S}^k_{t-1}\,\sigma^2_{\lambda|t,k}. \tag{26}$$

Substituting (24)–(26) into (23) yields:

$$\sigma^2_{S|t,k} = \sigma^2_{S|t-1,k}\Big(\sigma^2_{\lambda|t,k} + (\lambda^k_t)^2\Big) + \sigma^2_{\mu|t,k}\Big(\sigma^2_{\lambda|t,k} + (1-\lambda^k_t)^2\Big) + \sigma^2_{\lambda|t,k}\Big(L^k_t - \mu^k_t - \bar{S}^k_{t-1}\Big)^2. \tag{27}$$

Disagreement recursion. From (21), $(L^k_t - \mu^k_t - \bar{S}^k_{t-1}) = \Delta\bar{S}^k_t/(1-\lambda^k_t)$. Substituting into (27) gives the LS disagreement recursion:

$$\sigma^2_{S|t,k} = \sigma^2_{S|t-1,k}\Big(\sigma^2_{\lambda|t,k} + (\lambda^k_t)^2\Big) + \sigma^2_{\mu|t,k}\Big(\sigma^2_{\lambda|t,k} + (1-\lambda^k_t)^2\Big) + \sigma^2_{\lambda|t,k}\left[\frac{\Delta\bar{S}^k_t}{1-\lambda^k_t}\right]^2. \tag{28}$$

Equation (28) decomposes narrative disagreement into three components driven by heterogeneity in: (i) inherited priors (the term proportional to $\sigma^2_{S|t-1,k}$), (ii) weights on priors ($\sigma^2_{\lambda|t,k}$), and (iii) interpretation of public information ($\sigma^2_{\mu|t,k}$).

G.2 Estimation: Variance decomposition

Data objects used in estimation. For each topic k and month t we observe $S^k_{c,t}$ for countries $c = 1,\ldots,C$. We form sample analogs: $\bar{S}^k_t = C^{-1}\sum_c S^k_{c,t}$, $\Delta\bar{S}^k_t = \bar{S}^k_t - \bar{S}^k_{t-1}$, $\sigma^2_{S|t,k} =$

$C^{-1}\sum_c (S^k_{c,t} - \bar{S}^k_t)^2$. The LS estimation strategy does not require directly observing $(L^k_t - \mu^k_t)$ because it enters only via $\Delta\bar{S}^k_t$ in (28). We nonetheless interpret L^k_t as our constructed global topic information index.

G.2.1 Estimation of λ via the forecast-revision regression

Random-coefficient restriction and revision regression. Rewriting (4) as

$$S^k_{c,t} = \lambda^k_{c,t} S^k_{c,t-1} + \varepsilon^k_{c,t}, \qquad \varepsilon^k_{c,t} \equiv (1 - \lambda^k_{c,t})(L^k_t - \mu^k_{c,t}), \tag{29}$$

the fully heterogeneous $\{\lambda^k_{c,t}\}$ is not directly estimable without restrictions. Following LS, we impose a parsimonious random-coefficient structure:

$$\lambda^k_{c,t} = \lambda^k + v^k_c, \qquad \mathbb{E}_c[v^k_c] = 0, \qquad \mathrm{Var}_c(v^k_c) = \sigma^2_{\lambda|k}, \tag{30}$$

with v^k_c mutually independent across c and independent of $(S^k_{c,t-1}, \mu^k_{c,t})$ in the cross section. As in LS, we use the revision regression to avoid spurious dynamics:

$$\Delta S^k_{c,t} = \beta^k S^k_{c,t-1} + u^k_{c,t}, \qquad \beta^k \equiv \lambda^k - 1, \tag{31}$$

where $u^k_{c,t} = \varepsilon^k_{c,t} + v^k_c S^k_{c,t-1}$. Under (22) and (30), $u^k_{c,t}$ is uncorrelated with $S^k_{c,t-1}$.

Common correlated effects (CCE) augmentation. To allow for cross-country correlation from common aggregate shocks and topic-wide narratives, we follow LS and augment (31) with Pesaran (2006) CCE controls:

$$\Delta S^k_{c,t} = \beta^k S^k_{c,t-1} + \boldsymbol{c}^{k\prime} \boldsymbol{z}^k_t + \eta^k_{c,t}, \qquad \boldsymbol{z}^k_t \equiv \left(1,\ \bar{S}^k_{t-1},\ \Delta\bar{S}^k_t\right). \tag{32}$$

We estimate (32) separately by topic k using pooled OLS across all (c,t) observations available. We then recover

$$\widehat{\lambda}^k \equiv 1 + \widehat{\beta}^k, \qquad \widehat{\kappa}^k \equiv 1 - \widehat{\lambda}^k \quad \text{(weight on new public information)}. \tag{33}$$

For numerical stability in the variance-step below, and following LS, we impose an upper bound $\widehat{\lambda}^k \leq 0.95$.[26]

G.2.2 Estimation of σ^2_λ and σ^2_μ

Regression form of the variance recursion. Plugging $\widehat{\lambda}^k$ into (28) and rearranging yields the LS estimating equation (the analogue of LS Eq. (21) with calendar time in place of forecast horizon):

$$\sigma^2_{S|t,k} - (\widehat{\lambda}^k)^2 \sigma^2_{S|t-1,k} = \sigma^2_{\lambda|k}\sigma^2_{\mu|k} + \sigma^2_{\lambda|k}\left[\sigma^2_{S|t-1,k} + \left(\frac{\Delta \bar{S}^k_t}{1-\widehat{\lambda}^k}\right)^2\right] + \sigma^2_{\mu|k}(1-\widehat{\lambda}^k)^2 + e_{t,k}. \tag{34}$$

We impose *time invariance* of the heterogeneity parameters within each topic k, i.e. $(\sigma^2_{\lambda|t,k}, \sigma^2_{\mu|t,k}) = (\sigma^2_{\lambda|k}, \sigma^2_{\mu|k})$ for all t.

Define:

$$Y_{t,k} \equiv \sigma^2_{S|t,k} - (\widehat{\lambda}^k)^2 \sigma^2_{S|t-1,k}, \qquad X_{t,k} \equiv \sigma^2_{S|t-1,k} + \left(\frac{\Delta \bar{S}^k_t}{1-\widehat{\lambda}^k}\right)^2. \tag{35}$$

Under the time-invariance restriction, the two terms involving $\sigma^2_{\mu|k}$ in (34) can be collected into a single intercept:

$$Y_{t,k} = \sigma^2_{\lambda|k} X_{t,k} + \alpha_k + e_{t,k}, \qquad \alpha_k \equiv \sigma^2_{\mu|k}\left(\sigma^2_{\lambda|k} + (1-\widehat{\lambda}^k)^2\right). \tag{36}$$

Equation (36) is *linear* in the estimable regression coefficients $\left(\sigma^2_{\lambda|k}, \alpha_k\right)$; $\sigma^2_{\mu|k}$ is then recovered algebraically from α_k.

[26] This prevents the term $\left[\Delta \bar{S}^k_t/(1-\widehat{\lambda}^k)\right]^2$ from exploding when $\widehat{\lambda}^k$ is near one, as discussed in Lahiri and Sheng (2008) Section 4.2.

OLS estimation. We estimate (36) *separately for each topic k* by OLS over calendar months t (i.e., there is no pooling across topics, paralleling LS's horizon-by-horizon estimation).[27] Let $\widehat{\sigma}^2_{\lambda|k}$ denote the coefficient estimate on $X_{t,k}$ and $\widehat{\alpha}_k$ the estimated intercept.

Recovering $\widehat{\sigma}^2_{\mu|k}$. From (36),

$$\widehat{\sigma}^2_{\mu|k} \equiv \frac{\widehat{\alpha}_k}{\widehat{\sigma}^2_{\lambda|k} + (1-\widehat{\lambda}^k)^2}. \tag{37}$$

We obtain standard errors for $\widehat{\sigma}^2_{\mu|k}$ via the delta method using the HAC covariance matrix of $(\widehat{\alpha}_k, \widehat{\sigma}^2_{\lambda|k})$ (or via a moving-block bootstrap over months).

G.3 Counterfactual variance paths and channel attribution

Predicted disagreement path. For each topic k, given $\widehat{\lambda}^k$, $\widehat{\sigma}^2_{\lambda|k}$, and $\widehat{\sigma}^2_{\mu|k}$, we generate a fitted disagreement path by iterating (28) forward in calendar time:

$$\widehat{\sigma}^2_{S|t,k} = \widehat{\sigma}^2_{S|t-1,k}\left(\widehat{\sigma}^2_{\lambda|k} + (\widehat{\lambda}^k)^2\right) + \widehat{\sigma}^2_{\mu|k}\left(\widehat{\sigma}^2_{\lambda|k} + (1-\widehat{\lambda}^k)^2\right) + \widehat{\sigma}^2_{\lambda|k}\left[\frac{\Delta \bar{S}^k_t}{1-\widehat{\lambda}^k}\right]^2, \tag{38}$$

initialized at the first observed month t_0 with $\widehat{\sigma}^2_{S|t_0,k} = \sigma^2_{S|t_0,k}$.

Three channels. As in LS, define:

- P (priors / inherited disagreement): dispersion in initial priors propagated forward via λ;
- W (weights on priors): heterogeneity in $\lambda^k_{c,t}$ summarized by $\sigma^2_{\lambda|k}$;
- I (interpretation): heterogeneity in interpretation summarized by $\sigma^2_{\mu|k}$.

[27] LS pool *countries* within a fixed horizon h and use PCSE because the resulting TSCS residuals are contemporaneously correlated across countries. In our application, once we work *topic-by-topic*, (36) is a time-series regression in t, so we report heteroskedasticity- and autocorrelation-robust (HAC/Newey–West) standard errors (or, equivalently, a moving-block bootstrap over months).

Counterfactual variance paths. For any subset $A \subseteq \{P, W, I\}$, define the counterfactual path $v_{t,k}(A)$ as the forward iteration of (28) with channels not in A shut down:

- If $P \notin A$, set the initial variance to zero: $v_{t_0,k}(A) = 0$; otherwise set $v_{t_0,k}(A) = \sigma^2_{S|t_0,k}$.
- If $W \notin A$, set $\sigma^2_{\lambda|k} = 0$ everywhere it appears.
- If $I \notin A$, set $\sigma^2_{\mu|k} = 0$ everywhere it appears.

Thus, for $t > t_0$,

$$v_{t,k}(A) = v_{t-1,k}(A)\Big(\mathbb{1}\{W \in A\}\widehat{\sigma}^2_{\lambda|k} + (\widehat{\lambda}^k)^2\Big) + \mathbb{1}\{I \in A\}\widehat{\sigma}^2_{\mu|k}\Big(\mathbb{1}\{W \in A\}\widehat{\sigma}^2_{\lambda|k} + (1-\widehat{\lambda}^k)^2\Big) + \mathbb{1}\{W \in A\}\widehat{\sigma}^2_{\lambda|k}\left[\frac{\Delta \bar{S}^k_t}{1-\widehat{\lambda}^k}\right]^2. \tag{39}$$

LS-style ΔR^2 attribution (goodness-of-fit differences). To gauge the relative importance of the three disagreement channels, we follow the R^2-difference procedure in Lahiri and Sheng (2008). Fix a topic k and let observed disagreement be

$$D_{t,k} \equiv \sigma^2_{S|t,k}, \qquad t = t_0, \ldots, T_k,$$

where t_0 is the first month for which all inputs needed to iterate the disagreement recursion are available. For any subset of channels $A \subseteq \mathcal{N} \equiv \{P, W, I\}$, let $v_{t,k}(A)$ denote the *counterfactual predicted disagreement path* generated by iterating the Lahiri–Sheng recursion forward in time while shutting down channels not in A (as defined in (39)).

Step 1 (fit measure for a given subset A). For each subset $A \subseteq \mathcal{N}$, run the topic-specific time-series regression over months $t = t_0, \ldots, T_k$:

$$D_{t,k} \;=\; a^k_A \;+\; b^k_A\, v_{t,k}(A) \;+\; u_{t,k}(A), \tag{40}$$

and record its coefficient of determination, denoted $R^2_k(A)$. This $R^2_k(A)$ is a reduced-form

goodness-of-fit diagnostic that measures how well the path $v_{t,k}(A)$ tracks the time-series variation in observed disagreement $D_{t,k}$.

Step 2 (leave-one-out ΔR^2 for each channel). Let $R^2_{k,\text{full}} \equiv R^2_k(\mathcal{N})$ denote the fit when *all* channels are active. For each channel $j \in \{P, W, I\}$, define the leave-one-out fit

$$R^2_{k,-j} \;\equiv\; R^2_k(\mathcal{N} \setminus \{j\}),$$

and the LS incremental contribution of channel j as

$$\Delta R^2_{k,j} \;\equiv\; R^2_{k,\text{full}} - R^2_{k,-j}, \qquad j \in \{P, W, I\}. \tag{41}$$

Intuitively, $\Delta R^2_{k,j}$ measures the loss in explanatory power when channel j is shut down, holding fixed all other channels and the common inputs that enter the recursion (e.g., $\widehat{\lambda}^k$ and $\Delta\bar{S}^k_t$).

Step 3 (normalized shares). To report channel shares that sum to one for each topic k, define

$$\text{Share}^{\Delta R^2}_{k,j} \;\equiv\; \frac{\Delta R^2_{k,j}}{\Delta R^2_{k,P} + \Delta R^2_{k,W} + \Delta R^2_{k,I}}, \qquad j \in \{P, W, I\}. \tag{42}$$

Because $v_{t,k}(A)$ is generated by a nonlinear recursion and (40) is only a fit diagnostic, $\Delta R^2_{k,j}$ can occasionally be small or negative in finite samples. In that case, we follow the standard convention of setting negative $\Delta R^2_{k,j}$ to zero prior to normalization and interpret (42) as an approximate partition of explanatory power. Since $D_{t,k}$ and $v_{t,k}(A)$ are serially correlated, we quantify sampling uncertainty using a moving-block bootstrap over months: resample blocks of months, re-compute disagreement moments and re-generate all counterfactual paths $v_{t,k}(A)$, then recompute $\text{Share}^{\Delta R^2}_{k,j}$ in each bootstrap draw.

Shapley-value attribution (order-invariant decomposition of explanatory power). The LS ΔR^2 attribution is intuitive but order-dependent in nonlinear settings with interactions:

because the recursion includes interaction terms (notably involving the W channel) and the propagation term $v_{t-1,k}(A)$, the marginal contribution of a given channel can depend on which other channels are already active. We therefore also compute *Shapley values*, which allocate total explanatory power symmetrically by averaging marginal contributions over all possible channel orderings.

Value function. Fix a topic k and define the value of any subset $A \subseteq \mathcal{N}$ as

$$V_k(A) \;\equiv\; R^2_k(A),$$

where $R^2_k(A)$ is obtained from (40). (With an intercept in (40), $R^2_k(\emptyset) = 0$ because $v_{t,k}(\emptyset) \equiv 0$.)

General Shapley formula. Let $\mathcal{N} = \{P, W, I\}$ and $m \equiv |\mathcal{N}| = 3$. For channel $j \in \mathcal{N}$, the Shapley value is

$$\phi_k(j) = \sum_{A \subseteq \mathcal{N} \setminus \{j\}} \frac{|A|!\,(m - |A| - 1)!}{m!} \Big(V_k(A \cup \{j\}) - V_k(A) \Big). \tag{43}$$

The weights $\frac{|A|!\,(m-|A|-1)!}{m!}$ average the marginal contribution of j across all permutations of channel entry.

Closed-form expressions for three channels. When $m = 3$, (43) reduces to the following convenient closed forms (useful for coding):

$$\phi_k(P) = \frac{1}{3}\Big(R^2_k(P) - R^2_k(\emptyset)\Big) + \frac{1}{6}\Big(R^2_k(P,W) - R^2_k(W)\Big) + \frac{1}{6}\Big(R^2_k(P,I) - R^2_k(I)\Big) + \frac{1}{3}\Big(R^2_k(P,W,I) - R^2_k(W,I)\Big), \tag{44}$$

$$\phi_k(W) = \frac{1}{3}\Big(R^2_k(W) - R^2_k(\emptyset)\Big) + \frac{1}{6}\Big(R^2_k(P,W) - R^2_k(P)\Big) + \frac{1}{6}\Big(R^2_k(W,I) - R^2_k(I)\Big) + \frac{1}{3}\Big(R^2_k(P,W,I) - R^2_k(P,I)\Big), \tag{45}$$

$$\phi_k(I) = \frac{1}{3}\Big(R^2_k(I) - R^2_k(\emptyset)\Big) + \frac{1}{6}\Big(R^2_k(P,I) - R^2_k(P)\Big) + \frac{1}{6}\Big(R^2_k(W,I) - R^2_k(W)\Big) + \frac{1}{3}\Big(R^2_k(P,W,I) - R^2_k(P,W)\Big). \tag{46}$$

By Shapley efficiency, these satisfy

$$\phi_k(P) + \phi_k(W) + \phi_k(I) = R^2_k(P, W, I) - R^2_k(\emptyset) = R^2_{k,\text{full}}.$$

Normalized Shapley shares. We report topic-level Shapley shares as

$$\text{Share}^{\text{Shap}}_{k,j} \equiv \frac{\phi_k(j)}{\phi_k(P) + \phi_k(W) + \phi_k(I)} = \frac{\phi_k(j)}{R^2_{k,\text{full}}}, \qquad j \in \{P, W, I\}. \tag{47}$$

As with ΔR^2, finite-sample noise can occasionally yield small negative $\phi_k(j)$; when this occurs, we set negative $\phi_k(j)$ to zero prior to normalization and interpret (47) as an approximate decomposition. We also apply the same moving-block bootstrap over months.

H Correlations of National Belief Level Variables

Table A7: Correlations of National Belief Level Variables

	$TradeCN_c$	$TradeShare_c$	$AgreeCN_c$	BRI_c	$ContCN_c$	$DistCN_c$	$ComLangCN_c$	$RuleLaw_c$	$GovEff_c$	$RegQuality_c$	$ListedFirms_c$	$Turnover_c$	$RetCorrCN_c$
$TradeCN_c$	1.000												
$TradeShare_c$	-0.202	1.000											
$AgreeCN_c$	0.690	-0.540	1.000										
BRI_c	0.658	-0.031	0.591	1.000									
$ContCN_c$	-0.225	-0.109	0.289	-0.320	1.000								
$DistCN_c$	-0.626	0.059	-0.565	-0.490	-0.095	1.000							
$ComLangCN_c$	0.781	-0.131	0.462	0.354	-0.113	-0.205	1.000						
$RuleLaw_c$	-0.363	0.199	-0.693	-0.391	-0.347	0.478	0.040	1.000					
$GovEff_c$	-0.108	0.157	-0.527	-0.239	-0.421	0.335	0.300	0.953	1.000				
$RegQuality_c$	-0.244	0.109	-0.611	-0.283	-0.496	0.434	0.123	0.972	0.971	1.000			
$ListedFirms_c$	0.317	-0.062	0.133	0.109	-0.073	0.142	0.427	-0.116	-0.004	-0.058	1.000		
$Turnover_c$	-0.118	0.665	-0.545	-0.252	-0.036	-0.213	-0.222	0.191	0.134	0.088	0.086	1.000	
$RetCorrCN_c$	0.520	-0.143	0.357	0.212	0.190	-0.481	0.582	0.131	0.288	0.134	0.303	0.109	1.000

Notes: This table reports the pairwise correlations among the country-level explanatory variables used in equation (10).