Foreign Fundamentals Through Domestic Lenses: Cross-Country Interpretation Heterogeneity and International Capital Flows*

Isha Agarwal† Wentong Chen‡ Eswar Prasad§

July 2, 2026

Abstract

Common information about the same foreign economy is processed differently by investors in different source countries. We measure source-country interpretation heterogeneity using 1.5 million news articles about China from 16 economies and show that it is persistent, driven by within-topic framing rather than topic selection, and decomposable into inherited baseline stances and differential signal interpretation. A simple model of media-filtered belief formation predicts that interpretation heterogeneity has greater influence on portfolio allocation when investors have less destination-specific expertise and when destination fundamentals are more complex or opaque. Quasi-experimental variation from global non-China narrative shocks, interacted with predetermined topic composition, shows that exogenous shifts in source-country interpretation predict portfolio flows to China. These results identify source-country interpretation heterogeneity as a measurable, behaviorally relevant component of information frictions in international capital allocation.

JEL Classification Numbers: F30, G11, G15, D84
Keywords: interpretation heterogeneity, information frictions, cross-border capital flows, institutional investors, portfolio investment in China, media narratives, textual analysis, natural language processing

---

*The paper was previously circulated as "Beyond the Fundamentals: How Media-Driven Narratives Influence Cross-Border Capital Flows" and "Beyond the Fundamentals: Cross-Country Media Narrative Dispersion and International Capital Flows". The authors would like to thank Kenza Benhima (discussant), Tony Cookson, Ryan Israelsen (discussant), Shimon Kogan (discussant), Song Ma, Varun Sharma (discussant), Chenzi Xu (discussant), and seminar and conference participants at the Banque de France, UC Santa Cruz, University of Toronto, the Bank of England's 2026 BEAR conference 'The Interconnected World', MFA 2026, 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'. Rutvik Sitiraju and 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

Investors in different countries form persistently different interpretations of the same foreign economy. Consider how global media outlets covered China's stock market crash event in January 2016. The *Wall Street Journal* described "panic selling" and was pessimistic about broader impacts, while Singapore's *The Business Times* took a substantially more benign view.[1] These differences are not random: they are systematic, persistent, and, behaviorally relevant for international capital allocation. This paper measures source-country interpretation heterogeneity about the same foreign economy, decomposes its sources, and establishes that it is a quantitatively important component of information frictions in cross-border portfolio investment.

Understanding why investors in different countries form different beliefs about the same foreign economy is a central question in international finance. Allocating capital across borders requires processing large volumes of unfamiliar and often opaque information, complicated by language barriers, institutional differences, and policy uncertainty. Because investors rarely observe foreign economies directly, domestic news media play a central role in shaping investors' beliefs through editorial gatekeeping — selecting which foreign developments to cover and how to frame them (Shoemaker and Vos, 2009; Nimark and Pitschner, 2019; Chahrour, Nimark and Pitschner, 2021). If media in different countries systematically portray the same foreign economy differently, investors across source countries may form different beliefs about a common destination, causing capital flows to depend not only on destination-country fundamentals but also on how those fundamentals are interpreted in each source country's domestic media.

Yet we know surprisingly little about whether source-country interpretation of the same foreign economy differs systematically across countries, and whether such differences affect international capital flows. The literature on media and financial markets has expanded rapidly but remains overwhelmingly domestic in scope (Tetlock, 2007; Engelberg and Parsons, 2011; Manela and Moreira, 2017; Calomiris and Mamaysky, 2019; Baker et al., 2021; Ahern and Peress, 2023; Bybee et al., 2024). A separate strand constructs country-level or firm-level measures of uncertainty, risk, and sentiment from news text (Baker et al., 2016; Hassan et al., 2019; Calomiris and Mamaysky, 2019; Handley and Li, 2020; Ahir et al., 2022; Caldara and Iacoviello, 2022; Hassan et al., 2024; Arteaga-Garavito et al., 2024; van Binsbergen et al., 2024; Hoberg and Manela, 2025). Neither strand examines whether investors in different countries are exposed to systematically different interpretations of the

[1] See "China's Stocks Are Hit by 'Panic Selling,'" *Wall Street Journal*, January 5, 2016, and "Stock Market Impact on China's Economy Limited," *The Business Times*, January 25, 2016.

same destination economy, or whether such interpretation differences shape cross-border capital allocations. Our paper fills this gap.

We focus on China as it provides a setting with pronounced information frictions and is an increasingly important destination for global institutional investors. Investors face well-known challenges in accessing timely and reliable data on China, exacerbated by opacity in policymaking and periodic disruptions 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, source-country interpretation heterogeneity is likely to be both large and consequential.

Framework. We first develop a signal-extraction framework to discipline the empirical analysis. Investors from different source countries observe the same destination-country fundamentals but process them through different public information environments. The framework has three nested layers. In the public layer, each country's aggregate interpretation index decomposes into two components: the share of coverage devoted to each topic, capturing the extensive margin of attention allocation, and the interpretation within each topic, capturing the intensive margin of framing. Within each topic, a source-country interpretation index evolves as a precision-weighted combination of inherited interpretation and a bias-adjusted common signal. An interpretation wedge captures persistent cross-country differences in how the same destination-country news is translated into beliefs, while an updating weight captures heterogeneity in how strongly each country revises its interpretation in response to new information. In the private layer, each investor holds an accumulated private-information sufficient statistic that combines inherited private knowledge with new destination-specific signals, where private precision is increasing in investor expertise and decreasing in destination complexity. In the portfolio layer, investors combine source-country public interpretation and accumulated private information via Bayesian precision weighting, so that the weight on public interpretation is decreasing in private precision.

Cross-country disagreement in source-country interpretation operates through four nested channels. At the extensive margin, countries differ in their topic coverage shares. At the intensive margin, countries differ in their within-topic interpretation, which further decomposes into two components: belief-level differences, comprising heterogeneity in inherited baseline priors and persistent wedges in how countries map common signals into beliefs, and updating heterogeneity, reflecting differences in updating intensity that

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

are amplified during periods of large common signal revisions.

The framework delivers three empirical predictions. First, more favorable source-country interpretation raises destination allocations through the expected return channel. Second, investors with greater accumulated private information about the destination rely less on source-country public interpretation in forming portfolio allocations. Third, reliance on source-country interpretation increases when destination fundamentals are more complex or opaque, as higher uncertainty raises the cost of private information acquisition for all investors.

Measurement. Using modern natural language processing techniques, including transformer based sentiment models and large language models, we construct cross-country indices of median interpretation about China from more than one million articles published in 39 newspapers across 16 economies between 2007 and 2022. We construct two types of indices at the aggregate level and separately for each of twelve distinguishable topic categories using an LLM-based classifier. The attention index captures the extensive margin of media exposure at two levels: at the aggregate level, it measures the share of all news coverage allocated to China; at the topic level, it measures the share of China-related coverage devoted to each of twelve topic categories, capturing which aspects of China each country chooses to emphasize. The interpretation index captures the intensive margin of media framing: at the aggregate level, it measures the overall favorability of China-related coverage; at the topic level, it measures how favorably or adversely each country frames each topic category. We also develop a unified multilingual framework based on language-specific FinBERT-style models, enabling comparable measurement across domestic media environments without relying on translation or English-only reporting.

Characterizing Interpretation Heterogeneity. We begin by documenting substantial and persistent cross-country dispersion in median interpretation about China. Some countries consistently portray China negatively, while others adopt markedly more positive portrayals. To understand the sources of this heterogeneity, we develop a two-step decomposition. First, following the conceptual framework, we decompose aggregate interpretation into topic coverage and within-topic interpretation. We construct two counterfactual interpretation indices: one that holds topic coverage fixed across investor countries while allowing within-topic interpretation to vary, and another that holds within-topic interpretation fixed while allowing topic coverage to vary. We find that nearly all cross-country variation is driven by within-topic interpretation rather than topic selection: countries largely cover the same topics but evaluate them differently. This result rules

out the hypothesis that interpretation heterogeneity is simply a consequence of different editorial topic choices. Second, we decompose within-topic interpretation heterogeneity into two components. Belief-level heterogeneity comprises persistent country-specific baseline priors and interpretation wedges, reflecting how differently countries map the same common signal into beliefs. Updating heterogeneity captures cross-country differences in how strongly countries revise their interpretation in response to new information. Belief-level heterogeneity accounts for the majority of cross-country disagreement on average, while updating heterogeneity is relatively more important in topics with frequent and measurable information flows, such as economic growth and financial markets. This pattern suggests that cross-country interpretation differences reflect persistent narrative lenses rather than differences in information processing speed.

Portfolio Flows. We then investigate whether source-country interpretation heterogeneity influences cross-border capital flows to China. Using quarterly fund-level data from Morningstar for approximately 20,000 open-ended mutual funds domiciled in 16 economies, we find that domestic median interpretation of China has a significant and economically meaningful association with portfolio allocation decisions, even after controlling for China's macroeconomic and financial fundamentals and including fund and quarter fixed effects. Consistent with Prediction 1, a one-standard-deviation increase in the interpretation index is associated with a 3.8% annualized increase in investment flows into China. This result is robust to alternative sentiment measures including bag-of-words methods and multilingual FinBERT indices.

The central identification challenge is to disentangle source-country interpretation shifts from changes reflecting China's macroeconomic or financial fundamentals. We address this using two complementary approaches. First, we exploit the Arab Spring as a plausibly exogenous global shock to interpretation of China. Although China's economic fundamentals were unchanged, the ensuing shift in global political narratives led media outlets to reinterpret China through a geopolitical risk lens. We show that countries whose media environments are more oriented toward geopolitical reporting experience a pronounced deterioration in interpretation of China following the Arab Spring, despite no corresponding differential change in attention to China. These interpretation 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 rather than changes in destination fundamentals. Second, we construct a shift-share instrument for domestic Chinan interpretation using non-China-related coverage in the *Wall Street Journal*, interacted with predetermined

country-level topic composition. This design exploits exogenous variation in the global narrative environment while holding fixed countries' topic exposure shares, isolating variation in source-country interpretation that arises from differential exposure to global narrative fluctuations rather than from contemporaneous China-specific economic developments. Across both designs, source-country interpretation continues to predict portfolio flows to China, consistent with Prediction 1.

Mechanism. We provide direct evidence for the information-friction mechanism by testing the cross-sectional and time-series predictions of the framework. Consistent with Prediction 2, investors with greater private information about China, proxied by earlier China investment entry, higher China asset share, accumulated information stock, and geographic proximity, respond significantly less to source-country interpretation. Consistent with Prediction 3, allocation sensitivity to interpretation is higher during periods of elevated market uncertainty. These heterogeneity patterns are not consistent with a pure prior-reinforcement story, under which all investors within a source country would respond equally to domestic media. The attenuation with private information suggests that source-country interpretation carries independent informational content beyond investor priors.

Belief-Level versus Updating Heterogeneity in Flows. We further isolate the contribution of belief-level versus updating heterogeneity to portfolio flows by constructing counterfactual interpretation indices that hold one component fixed while allowing the other to vary. The belief-level counterfactual consistently exhibits stronger explanatory power for cross-border flows than the updating counterfactual. This pattern is consistent with Proposition 1: persistent interpretation differences, rather than heterogeneity in short-run updating, account for the majority of interpretation-driven capital flows.

Related Literature. Our paper contributes to three strands of literature. First, we contribute to the literature on news media and financial market outcomes, which remains overwhelmingly domestic in scope.[3] We are the first to construct cross-country media-based interpretation indices about the *same destination economy* using newspapers from *multiple investor countries*, and to show that interpretation heterogeneity about the same economy shapes international capital flows. Second, we contribute to the text-as-data

[3]Existing research has 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; Goldman, Gupta and Israelsen, 2024; Bybee, Kelly, Manela and Xiu, 2024; Hu, 2024; Macaulay and Song, 2024; Kaminski, Macaulay and Song, 2026).

literature and to the literature on belief dispersion (Dumas et al., 2017; Cookson et al., 2020; Cookson and Niessner, 2020; Cookson et al., 2023, 2024; Goldman et al., 2024). We bridge these literatures by introducing a media-based, cross-country measure of interpretation disagreement about the same foreign economy, decomposing it into topic selection and within-topic interpretation components, and linking it to portfolio flows. Our multilingual sentiment framework enables comparable measurement across domestic media environments without relying on translation. Third, we contribute to the literature on determinants of cross-border capital flows.[4] Existing studies emphasize information frictions as a key driver of international investment (Brennan and Cao, 1997; Ahearne, Griever and Warnock, 2004; Gelos and Wei, 2005; Portes and Rey, 2005; Van Nieuwerburgh and Veldkamp, 2009; Andrade and Chhaochharia, 2010; Karolyi, Ng and Prasad, 2020; Levchenko and Pandalai-Nayar, 2020; Benhima and Cordonier, 2022), but typically treat these frictions as model wedges or proxy them with geographic, institutional, or technological distance. We provide a direct, investor-facing measure of the interpretation environment, capturing not only information availability but also how common information is framed across source countries. The results show that source-country interpretation heterogeneity is a measurable and behaviorally relevant component of information frictions in international capital allocation.

2 A Framework for Source-Country Interpretation Heterogeneity

We develop a signal-extraction framework that disciplines the empirical analysis. Investors from different source countries face the same underlying destination country fundamentals, but those fundamentals are filtered through different public information environments before entering portfolio choice, generating heterogeneous capital flows across source countries. Private information further modulates how much each investor relies on the public environment.

[4]Previous studies have highlighted various factors affecting cross-border flows, including quality of governance (Leuz, Lins and Warnock, 2009), currency and foreign returns (Froot and Ramadorai, 2005; Camanho, Hau and Rey, 2022), 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), tax haven status (Coppola, Maggiori, Neiman and Schreger, 2021), investor experience and private information (Albuquerque, Bauer and Schneider, 2009; Malmendier, Pouzo and Vanasco, 2020), economic uncertainty (Alok, Javadekar, Kumar and Wermers, 2022), and broader barriers to international capital allocation such as taxes and political risk (Pellegrino, Spolaore and Wacziarg, 2025). 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).

2.1 Environment and Public Interpretation

There is a destination economy, and a set of investor source countries indexed by c. Time t is discrete. Let θ_t^k denote the payoff-relevant latent state of the destination country in topic k. Investors do not directly observe θ_t^k and instead infer it by combining public and private information.

The source-country media environment serves as the primary public information channel through which investors learn about the destination country. Media outlets in each country apply a country-specific interpretive lens when translating common destination-country news into local coverage. Following Lahiri and Sheng (2008), let $S_{c,t}^k$ denote the source-country median interpretation index capturing country c's narrative stance toward the destination country on topic k, interpreted as a proxy for the posterior mean belief about θ_t^k.

Country c enters period t with a Gaussian predictive prior:

$$\theta_t^k \,\big|\, \mathcal{I}_{c,t-1}^k \sim \mathcal{N}\Big(S_{c,t-1}^k,\, (\pi_{c,t|t-1}^k)^{-1}\Big),$$

where $\mathcal{I}_{c,t-1}^k$ denotes the source-country public information environment available at the end of period $t-1$, $S_{c,t-1}^k$ is the predictive prior mean, inherited from last period's posterior, and $\pi_{c,t|t-1}^k$ is the predictive prior precision for θ_t^k, which incorporates any depreciation of previously acquired information as destination fundamentals evolve between $t-1$ and t.[5]

All countries observe a common public signal L_t^k about θ_t^k but interpret it differently. Country c's signal estimate based only on the new public signal is

$$Y_{c,t}^k = L_t^k - \varepsilon_{c,t}^k, \qquad \varepsilon_{c,t}^k \sim \mathcal{N}\Big(\mu_{c,t}^k,\, (b_{c,t}^k)^{-1}\Big),$$

where $\mu_{c,t}^k$ is an *interpretation wedge*, 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. Importantly, individual investors observe the source-country public interpretation $S_{c,t}^k$ as a single object, the realized narrative stance of their local media environment, rather than observing L_t^k and $\mu_{c,t}^k$ separately. Investors therefore cannot debias their local media narrative even if they are fully rational Bayesian updaters.[6]

[5] We allow predictive prior precision to vary across countries, reflecting heterogeneity in information uncertainty arising from geographic distance, cultural proximity, and media infrastructure, consistent with standard assumptions in the forecasting literature (Lahiri and Liu, 2006; Lahiri and Sheng, 2008)

[6] This assumption is standard in models of media and belief formation (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006). It is also empirically plausible: a German fund manager reading local newspaper coverage of China observes a realized narrative, not a decomposition of that narrative into a common signal and a German-specific interpretive lens. The cross-country comparison that identifies the

Under Gaussian priors and signals, Bayes' rule implies the posterior mean belief $S^k_{c,t}$ is a convex combination of the predictive prior mean and the bias-adjusted signal:

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

where the *weight on the predictive prior* is

$$\lambda^k_{c,t} \equiv \frac{\pi^k_{c,t|t-1}}{\pi^k_{c,t|t-1} + b^k_{c,t}} \in (0,1),$$

and the posterior precision updates as:

$$\pi^k_{c,t} = \pi^k_{c,t|t-1} + b^k_{c,t}.$$

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

Lemma 1 (Topic Decomposition of Aggregate Interpretation)**.** *The aggregate interpretation index $S_{c,t}$ is a weighted average of topic-level interpretation:*

$$S_{c,t} = \sum_{k=1}^{K} w^k_{c,t} S^k_{c,t}$$

where $w^k_{c,t}$ is the share of media coverage devoted to topic k.

Cross-country variation in $S_{c,t}$ therefore arises either from differences in topic coverage $w^k_{c,t}$ or from differences in within-topic interpretation $S^k_{c,t}$.

2.2 Sources of Interpretation Heterogeneity

Having characterized how aggregate interpretation decomposes into topic coverage and within-topic interpretation, we now examine the sources of cross-country differences in within-topic interpretation itself.

Proposition 1 (Belief-Level Heterogeneity and Updating Heterogeneity as Sources of Disagreement)**.** *Cross-country differences in source-country interpretation $S^k_{c,t}$ arise through two distinct channels. First, persistent interpretation differences: countries with $\mu^k_c \neq \mu^k_{c'}$ diverge in how they map the same common signal into interpretation. Second, heterogeneous updating: countries with $\lambda^k_c \neq \lambda^k_{c'}$ place different weights on new common information versus inherited*

interpretation wedge requires simultaneous observation of coverage in multiple countries, which is precisely what our dataset provides but individual investors do not have.

interpretation, generating dynamic differences that are amplified during periods of large common signal revisions $\Delta\bar{S}_t^k$.[7] *Formally, the Bayesian updating model yields the following tractable recursion for cross-country variance or disagreement (Lahiri and Sheng, 2008):*

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

where

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

Proof. See Appendix B.1. ∎

Equation (2) 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; 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 (2) 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.

2.3 Private Information

Investor i in source country c holds a private information stock about the destination economy that evolves dynamically. Let $\mathcal{X}_{i,c,t}^k$ denote investor i's private information set about topic k, which includes current private research, past investment experience, analyst networks, and knowledge accumulated through prior destination-country exposure.

Define

$$P_{i,c,t}^k \equiv E[\theta_t^k \mid \mathcal{X}_{i,c,t}^k]$$

as the private-information sufficient statistic, with precision $\tau_{i,c,t}^k$. Analogous to the public layer, the investor's accumulated private belief is a weighted average of the inherited

[7] For example, heterogeneous priors and persistent disagreement are studied in Varian (1989); heterogeneous updating and its role in sustaining disagreement are documented in Cookson and Niessner (2020) and Dumas et al. (2017).

private belief and a new private signal $x^k_{i,c,t}$:

$$P^k_{i,c,t} = \varphi^k_{i,c,t} P^k_{i,c,t-1} + \left(1 - \varphi^k_{i,c,t}\right) x^k_{i,c,t},$$

where $\varphi^k_{i,c,t} \in (0,1)$ is the weight on accumulated past private belief, determined by the relative precision of inherited versus new private information. More derivations can be found in Appendix A.[8]

The precision of the accumulated private statistic reflects investor expertise and destination complexity:

$$\tau^k_{i,c,t} = \tau^P(\xi_i, \Omega_t), \qquad \frac{\partial \tau^P}{\partial \xi_i} > 0, \quad \frac{\partial \tau^P}{\partial \Omega_t} < 0,$$

where ξ_i denotes investor expertise and Ω_t denotes destination complexity.[9]

Treating $S^k_{c,t}$ as the source-country public interpretation sufficient statistic and $P^k_{i,c,t}$ as the investor's private sufficient statistic, Bayesian updating yields the portfolio belief:

$$\widehat{\theta}^k_{i,c,t} = \omega_{i,c,t} P^k_{i,c,t} + (1 - \omega_{i,c,t}) S^k_{c,t}, \qquad \omega_{i,c,t} = \frac{\tau^k_{i,c,t}}{\tau^k_{i,c,t} + \pi^k_{c,t}}. \tag{3}$$

The weight on source-country public interpretation $(1 - \omega_{i,c,t})$ is decreasing in private precision $\tau^k_{i,c,t}$. Investors who have accumulated more destination-specific private knowledge carry higher $\tau^k_{i,c,t}$ and therefore exhibit lower sensitivity to local media sentiment $S^k_{c,t}$.[10]

2.4 Portfolio Choice

Let the destination country excess return follow:

$$r^d_{t+1} = \alpha_d + \sum_k \beta_k \theta^k_t + \varepsilon^d_{t+1}, \qquad \varepsilon^d_{t+1} \sim \mathcal{N}(0, \sigma^2_{\varepsilon,d,t}), \tag{4}$$

where α_d is a destination-specific intercept, $\beta_k > 0$ links topic-k fundamentals to returns, and ε^d_{t+1} is idiosyncratic return noise independent of θ^k_t and of all signals.[11]

[8] The precision weight $\varphi^k_{i,c,t} = \tau^k_{i,c,t-1}/(\tau^k_{i,c,t-1} + r^k_{i,c,t})$ follows from Bayesian updating, where $r^k_{i,c,t}$ is the precision of the new private signal.

[9] This reduced-form specification is consistent with rational inattention and costly information acquisition frameworks (Sims, 2003; Van Nieuwerburgh and Veldkamp, 2009, 2010; Kacperczyk et al., 2016).

[10] The role of accumulated private information and investor experience in international portfolio choice is documented in Brennan and Cao (1997), Albuquerque et al. (2007), and Malmendier et al. (2020).

[11] This return process is standard in the international portfolio literature (Brennan and Cao, 1997; Merton, 1987; Van Nieuwerburgh and Veldkamp, 2009). The topic loadings $\beta_k > 0$ capture the direction in which each topic-level fundamental affects destination returns.

Posterior expected return. Given the portfolio belief $\widehat{\theta}^k_{i,c,t}$ from equation (3), investor i's posterior expected excess return is:

$$\widehat{\mu}^d_{i,c,t} = E_{i,c,t}[r^d_{t+1}] = \alpha_d + \sum_k \beta_k \widehat{\theta}^k_{i,c,t} = \alpha_d + \sum_k \beta_k \left(\omega_{i,c,t} P^k_{i,c,t} + (1-\omega_{i,c,t}) S^k_{c,t}\right). \tag{5}$$

Posterior return variance. The investor's perceived return variance has two components: residual return risk and uncertainty about the latent destination state:

$$\widehat{\sigma}^{2,d}_{i,c,t} = \mathrm{Var}_{i,c,t}(r^d_{t+1}) = \sigma^2_{\varepsilon,d,t} + \boldsymbol{\beta}' V^{\theta}_{i,c,t} \boldsymbol{\beta}, \tag{6}$$

where $V^{\theta}_{i,c,t}$ is the investor's posterior covariance matrix over topic-level destination states. Assuming topics are independent, this simplifies to:

$$\widehat{\sigma}^{2,d}_{i,c,t} = \sigma^2_{\varepsilon,d,t} + \sum_k \beta^2_k v^k_{i,c,t}, \qquad v^k_{i,c,t} = \frac{1}{\pi^k_{c,t} + \tau^k_{i,c,t}}, \tag{7}$$

where $v^k_{i,c,t}$ is the posterior variance of θ^k_t given all available information.[12] Investors with higher accumulated private precision $\tau^k_{i,c,t}$ or higher public precision $\pi^k_{c,t}$ face lower posterior uncertainty about destination fundamentals and therefore lower perceived return variance.

Portfolio demand. A mean-variance investor with risk aversion γ_i chooses destination exposure:

$$\max_{a^d_{i,c,t}} \left\{ a^d_{i,c,t} \widehat{\mu}^d_{i,c,t} - \frac{\gamma_i}{2} (a^d_{i,c,t})^2 \widehat{\sigma}^{2,d}_{i,c,t} \right\}.$$

The first-order condition is:

$$a^{d,*}_{i,c,t} = \frac{\widehat{\mu}^d_{i,c,t}}{\gamma_i \widehat{\sigma}^{2,d}_{i,c,t}}.$$

[12] Topic independence is maintained for tractability, following Van Nieuwerburgh and Veldkamp (2009), so cross-topic covariance terms in $V^{\theta}_{i,c,t}$ are set to zero.

Substituting the posterior expected return and variance[13]:

$$a_{i,c,t}^{d,*} = \frac{\alpha_d + \sum_k \beta_k \left(\omega_{i,c,t} P_{i,c,t}^k + (1-\omega_{i,c,t}) S_{c,t}^k\right)}{\gamma_i \left(\sigma_{\varepsilon,d,t}^2 + \sum_k \beta_k^2 v_{i,c,t}^k\right)}. \tag{8}$$

2.5 Empirical Predictions

Prediction 1 (Interpretation Predicts Destination Allocation)**.** *A more favorable source-country interpretation environment raises destination allocations. From equation (8):*

$$\frac{\partial a_{i,c,t}^{d,*}}{\partial S_{c,t}^k} = \frac{\beta_k (1-\omega_{i,c,t})}{\gamma_i \widehat{\sigma}_{i,c,t}^{2,d}} > 0.$$

The sensitivity is increasing in the topic loading β_k and decreasing in risk aversion γ_i and perceived return variance $\widehat{\sigma}_{i,c,t}^{2,d}$.

Prediction 2 (Expertise Attenuates Reliance on Public Interpretation)**.** *Since $\partial \tau^P / \partial \xi_i > 0$, investor expertise raises accumulated private precision $\tau_{i,c,t}^k$, reducing the weight on source-country public interpretation:*

$$\frac{\partial (1-\omega_{i,c,t})}{\partial \tau_{i,c,t}^k} = -\frac{\pi_{c,t}^k}{(\pi_{c,t}^k + \tau_{i,c,t}^k)^2} < 0.$$

Define $h_k \equiv \pi_{c,t}^k + \tau_{i,c,t}^k$ and $D \equiv \sigma_{\varepsilon,d,t}^2 + \sum_j \beta_j^2 / h_j$. The cross-partial of allocation sensitivity with respect to investor expertise is:

$$\frac{\partial^2 a_{i,c,t}^{d,*}}{\partial S_{c,t}^k \partial \xi_i} = -\frac{\beta_k \pi^k}{\gamma_i h_k^2 D^2} \left(\sigma_{\varepsilon,d,t}^2 + \sum_{j \neq k} \frac{\beta_j^2}{h_j}\right) \frac{\partial \tau_{i.c.t}^k}{\partial \xi_i} < 0$$

if and only if:

$$\sigma_{\varepsilon,d,t}^2 + \sum_{j \neq k} \frac{\beta_j^2}{h_j} > 0.$$

[13] Source-country public interpretation also affects perceived return variance through negative framing and risk content, generating asymmetric allocation responses. Following Soroka (2006) and Epstein and Schneider (2008), we augment the perceived return variance with adverse interpretation intensity and risk content terms, where negative interpretation operates through both the expected return and variance channels simultaneously while positive interpretation operates only through the expected return channel. This generates two additional predictions: adverse interpretations have larger allocation effects than favorable interpretations of equal magnitude, and heightened risk content in source-country interpretation reduces destination exposure through a distinct second-moment channel. Full derivations and empirical results are provided in Appendix D and Appendix M.

This condition holds whenever there is residual payoff risk or more than one payoff-relevant topic, and fails only in the knife-edge case of a single topic with no residual return risk. Investors with greater destination-specific expertise therefore exhibit unambiguously lower allocation sensitivity to source-country public interpretation $S^k_{c,t}$.

Proof. See Appendix C.1. ■

Prediction 3 (Destination Complexity Amplifies Reliance on Public Interpretation)**.** *Holding public precision $\pi^k_{c,t}$ fixed, greater destination complexity Ω_t reduces accumulated private precision $\tau^k_{i,c,t}$, raising the weight on source-country public interpretation:*

$$\left.\frac{\partial(1-\omega_{i,c,t})}{\partial\Omega_t}\right|_{\pi^k_{c,t}} = \frac{\pi^k_{c,t}}{(\pi^k_{c,t}+\tau^k_{i,c,t})^2}\cdot\left(-\frac{\partial\tau^k_{i,c,t}}{\partial\Omega_t}\right) > 0.$$

Using the same shorthand h_k and D as in Prediction 2, the cross-partial of allocation sensitivity with respect to destination complexity is:

$$\left.\frac{\partial^2 a^{d,*}_{i,c,t}}{\partial S^k_{c,t}\,\partial\Omega_t}\right|_{\pi^k_{c,t}} = -\frac{\beta_k\pi^k}{\gamma_i h_k^2 D^2}\left(\sigma^2_{\varepsilon,d,t}+\sum_{j\neq k}\frac{\beta_j^2}{h_j}\right)\frac{\partial\tau^k_{i,c,t}}{\partial\Omega_t} > 0$$

if and only if:

$$\sigma^2_{\varepsilon,d,t}+\sum_{j\neq k}\frac{\beta_j^2}{h_j} > 0.$$

Under the same near-universal regularity condition as Prediction 2, higher destination complexity unambiguously raises allocation sensitivity to source-country public interpretation. When destination fundamentals are more complex or opaque, acquiring precise private information becomes more costly for all investors, reducing equilibrium $\tau^k_{i,c,t}$ and increasing reliance on source-country public interpretation (Sims, 2003; Van Nieuwerburgh and Veldkamp, 2009).[14]

Proof. See Appendix C.2. ■

3 Measuring Source-Country Interpretation Heterogeneity

We now construct empirical counterparts of $S^k_{c,t}$, $w^k_{c,t}$, and μ^k_c introduced in Section 2. We first describe our methodology for quantifying media narratives about China using

[14]This prediction holds conditional on public precision $\pi^k_{c,t}$ remaining unchanged as Ω_t varies. If destination complexity also degrades public signal quality, the net effect is ambiguous. The maintained assumption is that Ω_t disproportionately raises private information acquisition costs relative to public signal degradation, which is plausible because source-country media aggregates across many articles and journalists, partially averaging out individual interpretation noise.

natural language processing (NLP) techniques. We begin by outlining the data collection process and then detail the construction of our narrative indices. For each 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.

3.1 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 E 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.[15] 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.

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.[16] Detailed information on the sample period and sample construction for each media outlet can be found in Table A1 in Appendix E. 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.

[15] Following Maggiori, Neiman and Schreger (2020) and Coppola, Maggiori, Neiman and Schreger (2021), we define the EMU as comprising 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.

[16] 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.

3.2 Index Construction

Using natural language processing methods, we construct four key indices: an attention index, an aggregate interpretation index, a topic-level share index, and a topic-level interpretation 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 interpretation index, $sen_{c,t}$, at the article level, which captures the tone of reporting about China.[17] 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 interpretation 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.

3.2.1 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. We then compute the country level attention index by averaging these scaled measures across all media outlets within each country.

Interpretation Index. We construct an interpretation 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.[18] 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

[17] We also construct attention and sentiment indices using non-English newspapers in the extended analysis.

[18] We use the term *interpretation index* throughout the paper to emphasize that media framing reflects country-specific interpretive lenses rather than objective sentiment. In figures and tables, this index is labeled *sentiment index* for brevity.

implementation is provided in Appendix H.

The model classifies text into three sentiment categories, positive, neutral, and negative, each accompanied by a confidence score.[19] 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. Consistent with journalistic conventions, 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.[20] This approach preserves the most sentiment-relevant content while maintaining comparability across a large corpus of documents and multiple language settings.

After extracting article-level sentiment labels, 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 interpretation 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 interpretation index constructed using the FinBERT method ($sen_{c,t}$). The correlation between this FinBERT-based index and the traditional bag-of-words–based interpretation index is 0.82, indicating a high degree of consistency across methods.[21]

Cross-Language interpretation 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

[19] Examples of article-level classifications are reported in Table A3.

[20] This truncation reflects a deliberate design choice, but it also addresses a mechanical constraint. BERT-based models, including FinBERT, operate under fixed input length constraints, and the average article in our dataset contains approximately 695 words. Our results are also robust to an alternative specification that uses the headline and the first 1,500 characters of each article.

[21] We further construct an interpretation index using bag-of-words–based method and a risk index. The methodology is reported in Appendix G.

positive, neutral, and negative, which allows us to incorporate non-English text in a manner that remains systematically comparable across countries.

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 (see Appendix H).

We include sentiment indices from seven additional newspapers in German, French, and Spanish, covering Germany, France, Spain, Switzerland, and Austria in our empirical analysis. 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.

3.2.2 Topic-level Index

We 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 as follows: (1) Trade & Supply Chain; (2) Domestic Economy & Growth; (3) Financial Markets, Banking & FX; (4) Real Estate, Debt & Financial Stability; (5) Technology, 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; and (12) Firms, Industries & Corporate Strategy.

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 (see Appendix I.1).[22]

[22]To reduce computational time and cost, we classify articles using only their titles rather than full text.

For the *Financial Times* (FT) archive, we follow the same taxonomy and classification procedure, but due to data usage restrictions, classification is implemented via a local deployment of the `mistral-7b-instruct.Q4_0` model using the `llama.cpp` inference framework.[23] 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 1 presents bigram word clouds for each topic, demonstrating that the classification method effectively captures the 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 interpretation 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

Random subsample tests confirm that title-based and full-text classifications produce near identical results.

[23] 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.

country c to obtain the monthly index.

Topic Interpretation Index. The topic interpretation 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.

3.3 Visualizing the Indices

Figure 2 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 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.

An examination of topic-level media narratives shows that 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 (see Figure A2). 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.

4 Characterizing Cross-Country Interpretation Heterogeneity

4.1 Cross-Country Interpretation Heterogeneity

Figure 3 highlights the substantial cross-country variation in aggregate attention and sentiment indices. Asian economies tend to devote greater attention to China-related news and hold more favorable views toward China than advanced economies.

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), typically receive higher coverage shares than other categories. 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: sentiment about 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. 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 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.

4.2 Topic Attention versus Within-Topic Sentiment Heterogeneity

Consistent with Lemma 1, aggregate interpretation heterogeneity $S_{c,t}$ decomposes into topic coverage $w_{c,t}^k$ and within-topic interpretation $S_{c,t}^k$. We ask which component dominates 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 interpretation 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 interpretation

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 $s_{c,t}^{k}$ 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 interpretation index. Formally, let $\text{Var}_{t}(\cdot)$ denote the cross-country variance at a given time t. We compute:

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

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

$$\text{Residual}_t = 1 - \text{Share}_t^{\text{att}} - \text{Share}_t^{\text{tone}}, \tag{9}$$

which captures cross-country variation arising from the interaction between topic attention and within-topic sentiment.

Figure 5 summarizes the results of the variance decomposition. 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 other words, 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.

4.3 Heterogeneous Belief Levels versus Heterogeneous Updating

Why do narratives diverge within the same topic? Consistent with Proposition 1, within-topic disagreement in $S_{c,t}^k$ can be decomposed into belief-level heterogeneity and updating heterogeneity. We estimate the relative importance of each channel.

4.3.1 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 (1). 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:

$$\widehat{L}_{-c,t}^k \equiv \text{PC1}_t\Big(\{S_{c',t}^k\}_{c' \neq c}\Big), \tag{10}$$

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 $\widehat{L}^k_{-c,t}$ so that it has positive correlation with the cross-country mean $\bar{S}^k_t$.

Country-specific interpretation wedges and updating weights. Imposing time-invariant country–topic parameters $(\lambda^k_{c,t}, \mu^k_{c,t}) = (\lambda^k_c, \mu^k_c)$ and replacing L^k_t with $\widehat{L}^k_{-c,t}$ in (1), subtracting $\widehat{L}^k_{-c,t}$ from both sides yields:

$$\begin{aligned} S^k_{c,t} - \widehat{L}^k_{-c,t} &= \lambda^k_c\left(S^k_{c,t-1} - \widehat{L}^k_{-c,t}\right) - (1-\lambda^k_c)\mu^k_c + u^k_{c,t} \\ &\equiv \delta^k_c + \lambda^k_c\left(S^k_{c,t-1} - \widehat{L}^k_{-c,t}\right) + u^k_{c,t}, \end{aligned} \tag{11}$$

where $\delta^k_c \equiv -(1-\lambda^k_c)\mu^k_c$ is a country–topic intercept and $u^k_{c,t}$ collects measurement error in $\widehat{L}^k_{-c,t}$ and any deviations from time-invariant (λ^k_c, μ^k_c).

Estimation. We estimate (11) separately by topic k on the pooled (c,t) panel, allowing for country-specific λ^k_c and δ^k_c via interactions with country indicators. Standard errors are clustered at the country level, and we impose $\widehat{\lambda}^k_c \in [0, 0.95]$ for numerical stability. We then recover the implied interpretation wedge and updating intensity:

$$\widehat{\mu}^k_c \equiv -\frac{\widehat{\delta}^k_c}{1-\widehat{\lambda}^k_c}, \qquad \widehat{\kappa}^k_c \equiv 1 - \widehat{\lambda}^k_c. \tag{12}$$

Visualization: belief Level heterogeneity and updating heterogeneity. Figure 6 visualizes the time-invariant country–topic parameters implied by the Bayesian updating rule in (1). Panel (a) reports the estimated interpretation wedge $-\widehat{\mu}^k_c$ (higher values indicate a *more favorable* mapping from the common topic signal into sentiment), while Panel (b) reports the updating intensity $\widehat{\kappa}^k_c \equiv 1 - \widehat{\lambda}^k_c$, 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*: $-\widehat{\mu}^k_c$ exhibits large and easily visible dispersion across countries within most topics, whereas $\widehat{\kappa}^k_c$ 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.[24]

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.

4.3.2 Variance Decomposition: Belief-Level and Updating Heterogeneity

Empirical implementation. Following Proposition 1, 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 associated R^2. The marginal contribution of a given channel is measured by the change in explanatory power when that channel is removed. See Appendix B.1 for all derivations, identifying assumptions, and estimation details, and alternative specifications using an order-invariant Shapley decomposition.

Results. Figure 7 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

[24] See Appendix K for a correlation analysis between these national belief levels and country fundamentals related to China.

information. Belief-level components carry most of the variance across topics (on average $P{+}I \simeq 70\%$), 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 decomposition, 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., $P{+}I \simeq 79\%/87\%/78\%$), 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 Cross-Country Interpretation Heterogeneity and Portfolio Allocation

Prediction 1 implies that more favorable source-country interpretation predicts higher destination allocation. We test this using 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.

Motivated by Prediction 1, we estimate the impact of local media narratives on cross-

border flows of institutional investors using the following regression specification[25]:

$$Flow_{ic,t} = \beta_0 + \beta_1 Index_{c,t-1} + \beta_2 num_{c,t-1} + X'_{c,t-1}\Phi + \alpha_i + \alpha_t + \varepsilon_{ic,t} \tag{13}$$

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 flows 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.[26]

$Index_{c,t-1}$ denotes the media narrative based interpretation 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.[27] $X_{c,t-1}$ 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 ($Growthdiff_{c,t-1}$); the bilateral exchange rate 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 ($Volratio_{t-1}$).[28] 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

[25] In the theoretical framework, $a^{d,*}_{i,c,t}$ denotes the optimal portfolio share allocated to the destination country, which is the natural object in a mean-variance portfolio problem. In the empirical analysis, we follow the international portfolio literature and use net fund flows as the primary outcome variable, since portfolio shares are mechanically affected by valuation changes and are therefore a noisier measure of active portfolio reallocation decisions.

[26] Since it is hard to interpret valuation changes for derivative assets, we exclude all derivative holdings from our analysis. $Flow_{ic,t}$ has been winsorized at the 0.5% level in each tail.

[27] 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 13.

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

the fund level to adjust for potential serial correlation in the error term $\varepsilon_{ic,t}$.

Table 2 reports baseline estimates from alternative specifications of equation (13) using different combinations of control variables. Consistent with Prediction 1, 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 median interpretation 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.

5.1 Robustness Analysis

Bag-of-word Method. We further construct bag-of-word sentiment indices ($sen_{c,t}^{bag}$) using the method described in Appendix G, 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 3 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.

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 3 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 interpretation 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.[29]

The strong statistical significance and consistent signs across languages 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 Identification: Quasi-Experimental Evidence

The baseline results document a strong association between domestic media sentiment toward China and cross-border investment flows. A key identification challenge 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. We provide quasi-experimental evidence that exogenous shifts in source-country interpretation, holding destination fundamentals fixed, predict portfolio flows, consistent with the mechanism in Prediction 1.

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

First, we exploit 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-

[29] Domestic language newspapers tend to 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 language media markets. Multilingual sentiment indices therefore display lower cross country dispersion, even though their informational content remains highly relevant. 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.

oriented lens.[30] 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. 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.

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 search query approach 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.

Media Narrative Response to the Arab Spring. We define a post-event indicator $Post_t$ equal to one for all months from January 2011 onward, and estimate the following difference-in-differences specification at the country–month level:

$$index_{c,t} = \alpha + \beta\,(Post_t \times HighGeoPolitics_c) + \gamma_c + \delta_t + \varepsilon_{c,t}, \tag{14}$$

[30] 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.' ...*".

where $index_{c,t}$ denotes alternative country-level media narrative indices about China, including a China-related attention index and an aggregate interpretation 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 4 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.

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:

$$Flow_{ic,t} = \theta\,(Post_t \times HighGeoPolitics_c) + \alpha_i + \alpha_t + X'_{c,t-1}\Phi + \varepsilon_{ic,t}, \tag{15}$$

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 4 indicate 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.

6.2 A Shift–Share Method Using Non-China News

Next, we turn to 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 China-specific economic developments.

6.2.1 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 interpretation index can be equivalently constructed as a topic-share-weighted average of within-topic sentiment,

$$sen_{c,t} = \sum_{k=1}^{K} w_{c,k,t}\, sen_{c,t}^{(k)}, \tag{16}$$

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

$$\bar{w}_{c,k} = \frac{1}{|\mathcal{T}_0|} \sum_{t \in \mathcal{T}_0} w_{c,k,t}, \qquad \sum_{k=1}^{K} \bar{w}_{c,k} = 1. \tag{17}$$

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 3. 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,

$$Z_{c,t} = \sum_{k=1}^{K} \bar{w}_{c,k} \, G_{k,t}. \tag{18}$$

The instrument varies over time with global topic-level narrative shocks and varies across countries with persistent differences in topic attention.

6.2.2 Methodology

First Stage. At the first stage, we assess the relevance of the instrument by estimating the regression

$$sen_{c,t} = \pi Z_{c,t} + \alpha_c + \alpha_t + u_{c,t} \tag{19}$$

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 median interpretation 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

$$Flow_{ic,t} = \beta_0 + \beta_1 \widehat{sen}^{IV}_{c,t-1} + \beta_2 num_{c,t-1} + X'_{c,t-1}\Phi + \alpha_i + \alpha_t + \varepsilon_{ic,t} \tag{20}$$

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

Identification Assumptions. This identification 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. 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.

6.2.3 Effect on Capital Flows

The right-hand side of Table 5 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. The left-hand side 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 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. 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.

7 Mechanism Tests

To deepen our understanding of how domestic media narratives influence investment flows to China, we further explore whether media narratives alleviate frictions in information access or simplify the interpretation of available data. We leverage the rich heterogeneity

among open-ended funds investing in China to examine variations in their information sets and how this variation influences their investment response to changes in local media narratives about China. Additionally, we analyze how the impact of local media narratives differs across periods of high and low market volatility, given that interpreting market data could be more challenging during times of high volatility.

7.1 Expertise Attenuates Reliance on Public Interpretation

Prediction 2 implies that investors with greater expertise rely less on source-country interpretation. We proxy investor expertise ξ_i with four measures.

Initial Investment Dates. Institutional investors initiated their investments in China at different times. We infer the first investment date from the earliest recorded purchase of Chinese assets in each fund. When direct records are unavailable, we rely on the earliest date the fund held Chinese assets in its portfolio, as documented by Morningstar. We hypothesize that institutional investors who began investing in China earlier exhibit different sensitivities to media narratives compared to those who entered more recently. A longer investment tenure presumably correlates with a better understanding of Chinese market dynamics, potentially reducing reliance on media narratives when making investment decisions. We construct a dummy variable, $EarlyStart_i$, which equals one if a fund's initial investment date in China precedes the median initial investment date of all open-ended funds holding Chinese assets in our dataset, and zero otherwise.

Chinese Asset Share. The share of a fund's total assets invested in China might affect its reliance on local media narratives about China. On one hand, institutional investors with significant holdings in Chinese assets could be less influenced by media narratives about China due to the private information they access through their substantial investments. On the other hand, these investors may also have a higher demand for information about China, resulting in increased sensitivity to local media narratives about changes within the country. We construct a dummy variable, $HighShare_{i,t}$, which equals one if the Chinese asset holdings of fund i at quarter t exceed the median holdings of all open-ended funds with Chinese asset holdings in our dataset, and zero otherwise.

Stock of Information. We further explore the historical stock of information about China held by institutional investors by combining data on their initial investment start dates with their share of Chinese asset holdings. Intuitively, investors with earlier start dates and high shares of Chinese assets are likely to have accumulated extensive private information

about these assets, reflecting a high information stock and reduced reliance on local media narratives. We construct the dummy variable $HighInfoStock_{i,t}$, which is set to one if fund i's initial investment date in China predates the median among all open-ended funds holding Chinese assets, and if at quarter t, this fund's holdings of Chinese assets exceed the median holdings of all open-ended funds with Chinese assets in our dataset.

Geographical Proximity. Finally, we investigate whether investors from economies geographically closer to China exhibit distinct sensitivity to local media narratives about China. Intuitively, these investors are likely to have greater familiarity with China, attributable to geographical proximity, cultural similarities, or alignment in time zones. We construct a dummy variable $GeoClose_{i,t}$, which is set to one for funds domiciled in countries closer to China including Hong Kong, India, Malaysia, the Philippines, Singapore, South Korea, Taiwan, and Thailand.

To study how the different information sets of institutional investors affect their investment response to local media narratives about China, we estimate the following equation:

$$Flow_{ic,t} = \beta_0 + \beta_1 Index_{c,t-1} + \beta_2 Het_{i,t-1} + \beta_3 Index_{c,t-1} \times Het_{i,t-1} + \beta_4 num_{c,t-1} + \beta_5 X_{c,t-1} + \\ + \alpha_i + \alpha_t + \varepsilon_{ic,t} \quad (21)$$

where $Het_{i,t-1}$ is a dummy variable capturing fund-level heterogeneity along the four dimensions discussed above: $EarlyStart_i$, $HighShare_{i,t}$, $HighInfoStock_{i,t}$, or $GeoClose_{i,t}$. β_3 is our key coefficient of interest and measures the extent to which having better or private information about China dampens the impact of media narratives on investment flows to China.

Results. Table 8 reports the estimated results. Column (1) shows investors who began investing in China earlier reveal a 1.74% lower quarterly increase in investment for every one standard deviation increase in the local median interpretation index, suggesting that having an earlier start date in Chinese markets either increases the stock of private information that these investors have about China or enables them to better interpret market data, making them less reliant on local media narratives.

Column (2) indicates that institutional investors with a high share invested in Chinese assets see a 0.21% smaller increase in their investments for each one standard deviation increase in local median interpretation index in China. However, the effect is not significant. This could be because of two offsetting forces: investors with substantial Chinese asset holdings may rely less on local media due to more private information, or they may depend

more on local narratives for additional information due to their significant information needs.

Column (3) shows that investors with substantial stock of information about China experience a 1.08% smaller quarterly increase in their investments for a one standard deviation increase in the local median interpretation index.

Finally, as observed in column (4), investors geographically close to China experience a 1.90% smaller quarterly increase in their investments for each one standard deviation increase in local median interpretation index. These investors likely possess more information about China due to their proximity, cultural ties, or shared time zones and therefore, do not rely as much on local media.

These results suggest that local media narratives significantly impact cross-border investment flows into China via the information channel. Institutional investors who entered the Chinese market earlier, who hold substantial stocks of information about China, and who are from economies geographically close to China are likely to have access to alternative sources of private information about Chinese assets or can interpret Chinese market data better. This reduces their reliance on local media for news about China.

7.2 Destination Complexity Amplifies Reliance on Public Interpretation

Prediction 3 implies that greater destination complexity amplifies investor reliance on source-country public interpretation. We proxy destination complexity Ω_t using periods of elevated stock market volatility, on the grounds that high volatility makes destination fundamentals harder to interpret and raises the cost of private information acquisition. Specifically, we construct a dummy variable $HighVol_t$ equal to one if Chinese stock market volatility in that quarter exceeds the 75th percentile of all quarters in our sample, and zero otherwise.

Column (5) in Table 8 presents results with an interaction term between the narrative index and the dummy variable. After controlling for macroeconomic fundamentals, a one standard deviation increase in the interpretation index is associated with an additional 1.64% increase in quarterly investment flow into Chinese assets during high volatility periods, compared to other times. When market volatility is high, price signals become noisier and harder to interpret, leading investors to rely more heavily on local media narratives to contextualize economic and financial developments. At the same time, the positive interaction effect for the risk index suggests that, in turbulent periods, cross border investors may be more willing to take on risk, making the response to risk narratives less straightforward and highlighting the nuanced role of media narratives under stressed market conditions.

8 Belief Levels vs. Updating: Which Drives Flows?

Proposition 1 decomposes interpretation heterogeneity into belief-level ($\mathbf{P}+\mathbf{I}$) and updating ($\mathbf{W}$) components. We further document that cross-country media narrative dispersion is primarily driven by within-topic differences in sentiment, and that these within-topic differences are largely attributable to belief-level heterogeneity rather than updating heterogeneity across countries. We next construct belief-level and updating-driven media narrative wedge indices and test their relative importance for portfolio flows. Using the estimated parameters from equation (1), we recursively construct the full fitted interpretation index at the aggregate level:

$$\hat{s}^{full}_{c,t} = \hat{\lambda}_c \hat{s}^{full}_{c,t-1} + (1-\hat{\lambda}_c)(\hat{L}_t - \hat{\mu}_c).$$

where $\hat{L}_t$ is the estimated common signal, $\hat{\mu}_c$ is the estimated persistent interpretation wedge, and $\hat{\lambda}_c$ is estimated prior weight.

We further define a *belief-level counterfactual* index as:

$$Belief_{c,t} \equiv \hat{s}^{eqU}_{c,t} = \bar{\lambda}\hat{s}^{eqU}_{c,t-1} + (1-\bar{\lambda})(\hat{L}_t - \hat{\mu}_c), \qquad \bar{\lambda} = \bar{\hat{\lambda}}_c,$$

which equalizes the updating parameter across countries while preserving country-specific persistent interpretation wedges.

We define an *updating counterfactual* index as:

$$Updating_{c,t} \equiv \hat{s}^{eqB}_{c,t} = \hat{\lambda}_c \hat{s}^{eqB}_{c,t-1} + (1-\hat{\lambda}_c)(\hat{L}_t - \bar{\mu}), \qquad \hat{s}^{eqB}_{c,t_0} = \bar{s}_{t_0},$$

which sets the country-specific persistent interpretation wedge and initial belief level to the cross-country mean, while preserving heterogeneous updating intensities $\hat{\lambda}_c$.

To examine how these components affect capital flows, we replace the aggregate narrative index with the belief and updating wedges in the following specification:

$$Flow_{ic,t} = \beta_0 + \beta_1 Belief_{c,t-1} + \beta_2 Updating_{c,t-1} + \beta_3 num_{c,t-1} + X'_{c,t-1}\Phi + \alpha_i + \alpha_t + \varepsilon_{ic,t} \quad (22)$$

Table 7 reports the effects of belief-level and updating wedges on cross-border portfolio flows into China. We find that the belief-level wedge, capturing persistent cross-country differences in baseline narrative stance, has a statistically and economically significant effect on flows. A one-unit increase in the belief-level wedge is associated with a 6.79 % increase in quarterly portfolio allocation to Chinese assets. In contrast, the coefficient on the updating wedge, which captures heterogeneity in updating intensity conditional on

common information, is small and statistically insignificant across specifications.

9 Conclusion

This paper asks why investors in different countries form persistently different interpretations of the same foreign economy, and whether such differences matter for international capital allocation. Using more than one million newspaper articles from 16 economies, we construct source-country interpretation indices about China, decompose their sources, and link them to fund-level portfolio flows.

Our first set of findings characterizes the structure of interpretation heterogeneity. Cross-country dispersion in aggregate interpretation is driven almost entirely by within-topic framing rather than topic selection: countries largely cover the same aspects of China but evaluate them differently. Within-topic interpretation heterogeneity itself reflects persistent belief-level differences — comprising inherited baseline stances and systematic differences in how common signals are translated into beliefs — rather than heterogeneous updating intensity. These results suggest that source-country interpretation of foreign economies reflects slow-moving narrative lenses, not merely differences in information processing speed.

Our second set of findings establishes the behavioral relevance of interpretation heterogeneity. Domestic interpretation of China predicts institutional portfolio flows, even after controlling for macroeconomic and financial fundamentals and including fund and quarter fixed effects. This relationship holds across alternative sentiment measures, languages, and identification strategies. Consistent with the information-friction mechanism, the response to source-country interpretation is attenuated among investors with greater expertise about China, those with earlier entry, higher information stock, and greater geographic proximity, and amplified during periods of elevated uncertainty. The belief-level component of interpretation heterogeneity accounts for the majority of interpretation-driven flows, while updating heterogeneity plays a secondary role.

Our findings identify source-country interpretation heterogeneity as a measurable, persistent, and behaviorally relevant component of information frictions in international capital allocation. Disagreement in international finance is not solely the result of differential information access; it also reflects systematic differences in how common information is framed and interpreted by domestic information intermediaries.

Our results open several avenues for future research. Media-based measures of interpretation heterogeneity could be used to study cross-country differences in reactions to global shocks, policy announcements, and geopolitical events. Extending the analysis to

other destination economies would illuminate whether interpretation heterogeneity is more pronounced in environments characterized by greater opacity or policy uncertainty. Linking source-country interpretation to asset prices and risk premia across borders would further clarify the role of belief formation in global financial markets.

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.

Ahearne, Alan G, William L Griever, and Francis E Warnock (2004) "Information costs and home bias: an analysis of US holdings of foreign equities," *Journal of international economics*, 62 (2), 313–336.

Ahern, Kenneth R and Joel Peress (2023) *The role of media in financial decision-making*: Edward Elgar Publishing.

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.

Albuquerque, Rui, Gregory H. Bauer, and Martin Schneider (2007) "International Equity Flows and Returns: A Quantitative Equilibrium Approach," *Review of Economic Studies*, 74 (1), 1–30.

Albuquerque, Rui, Gregory H Bauer, and Martin Schneider (2009) "Global private information in international equity markets," *Journal of Financial Economics*, 94 (1), 18–46.

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.

Benhima, Kenza and Rachel Cordonier (2022) "News, sentiment and capital flows," *Journal of International Economics*, 137, 103621.

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.

Brennan, Michael J and H Henry Cao (1997) "International portfolio investment flows," *The Journal of Finance*, 52 (5), 1851–1880.

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.

Camanho, Nelson, Harald Hau, and Hélene Rey (2022) "Global portfolio rebalancing and exchange rates," *The Review of Financial Studies*, 35 (11), 5228–5274.

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," *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.

Dumas, Bernard, Karen K Lewis, and Emilio Osambela (2017) "Differences of opinion and international equity markets," *The Review of Financial Studies*, 30 (3), 750–800.

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

Epstein, Larry G. and Martin Schneider (2008) "Ambiguity, Information Quality, and Asset Pricing," *Journal of Finance*, 63 (1), 197–228.

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.

Gelos, R Gaston and Shang-Jin Wei (2005) "Transparency and international portfolio holdings," *The Journal of Finance*, 60 (6), 2987–3020.

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.

Goldman, Eitan, Nandini Gupta, and Ryan Israelsen (2024) "Political polarization in financial news," *Journal of Financial Economics*, 155, 103816.

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*.

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.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp (2016) "A rational theory of mutual funds' attention allocation," *Econometrica*, 84 (2), 571–626.

Kaminski, Alexa, Alistair Macaulay, and Wenting Song (2026) "Monetary Policy Narratives and the Transmission of Monetary Policy," Working paper.

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.

Levchenko, Andrei A and Nitya Pandalai-Nayar (2020) "TFP, news, and "sentiments": The international transmission of business cycles," *Journal of the European Economic Association*, 18 (1), 302–341.

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.

Macaulay, Alistair and Wenting Song (2024) "Narrative-Driven Fluctuations in Sentiment: Evidence Linking Traditional and Social Media," Working paper.

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

Malmendier, Ulrike, Demian Pouzo, and Victoria Vanasco (2020) "Investor experiences and financial market dynamics," *Journal of Financial Economics*, 136 (3), 597–622.

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.

Merton, Robert C. (1987) "A Simple Model of Capital Market Equilibrium with Incomplete Information," *Journal of Finance*, 42 (3), 483–510.

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.

Pellegrino, Bruno, Enrico Spolaore, and Romain Wacziarg (2025) "Barriers to global capital allocation," *The Quarterly Journal of Economics*, 140 (4), 3067–3131.

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.

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," *Journal of Finance*, 64 (3), 1187–1215.

——— (2010) "Information Acquisition and Under-Diversification," *Review of Economic Studies*, 77 (2), 779–805.

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

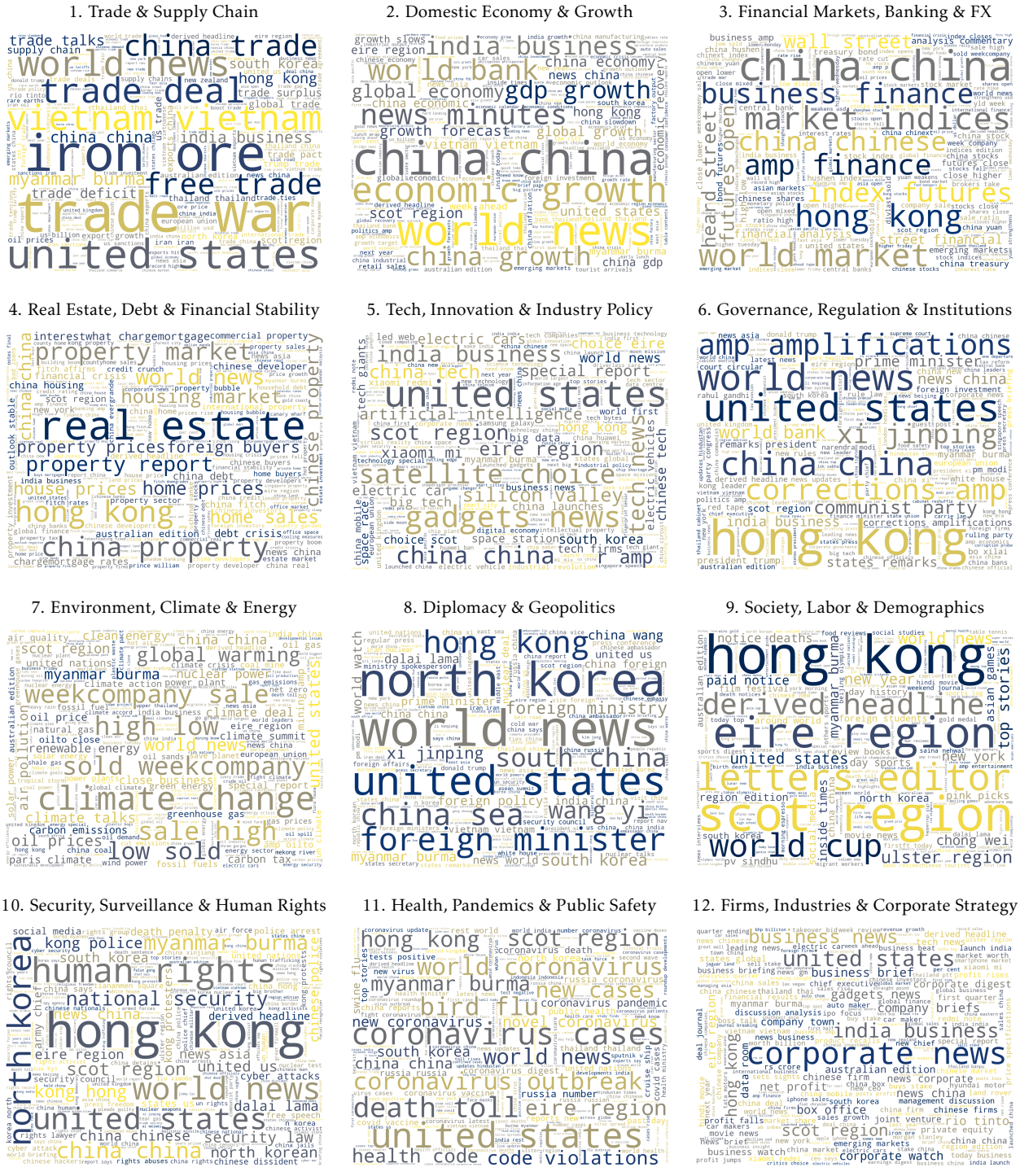
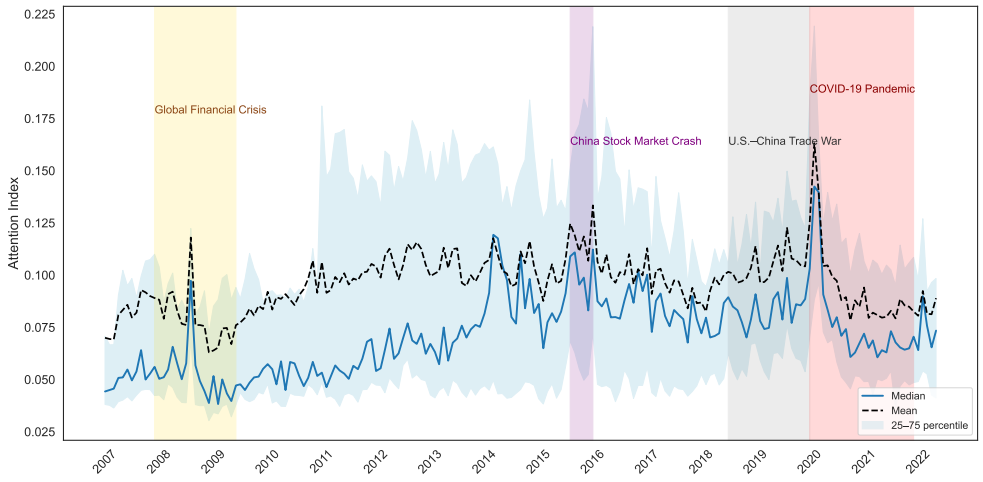


Figure 1: Bigram Word Clouds Across Topics

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

(a) Attention index



(b) Interpretation index

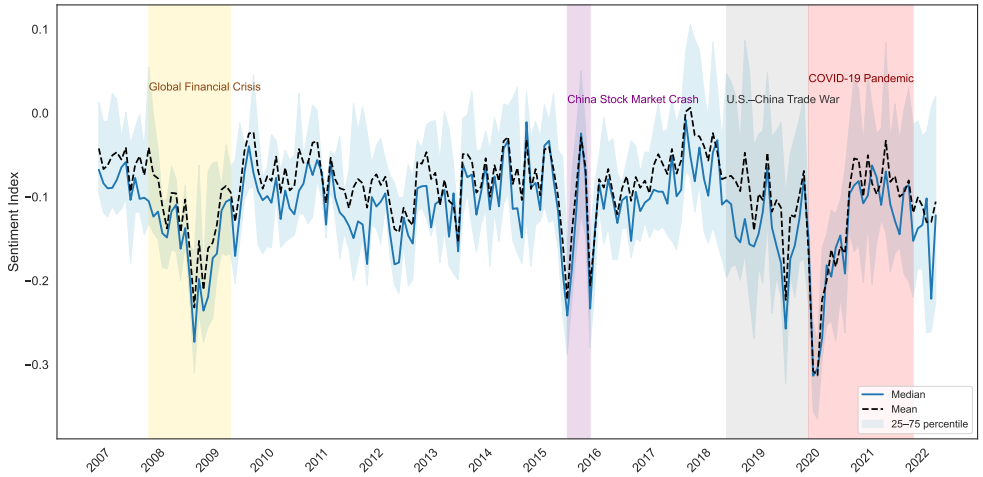
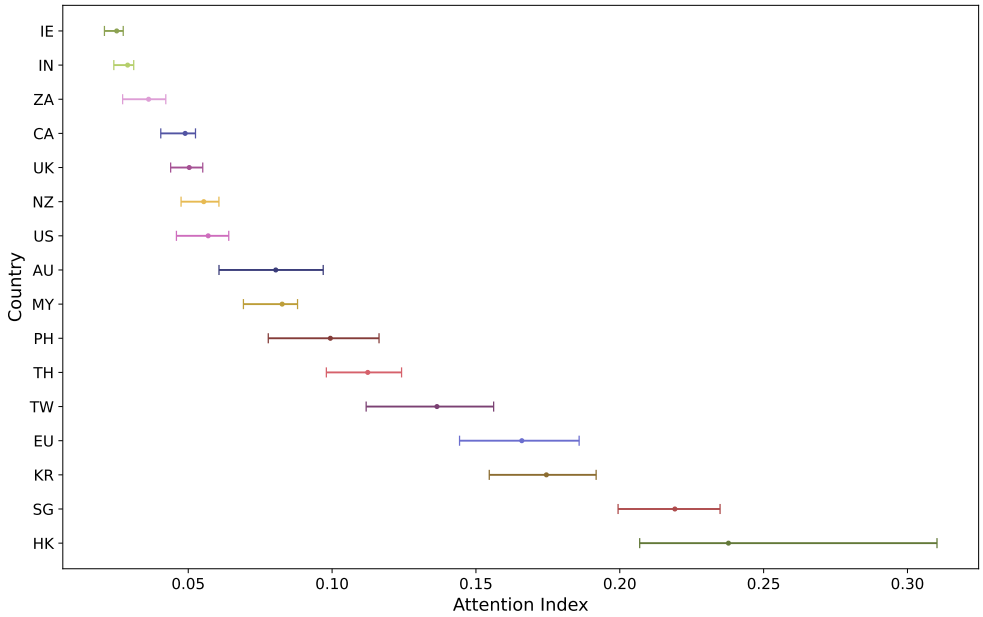


Figure 2: China Attention and Sentiment Indices

Notes: This figure plots the cross-country distribution of (a) the attention index and (b) the interpretation index for China over the period 2007–2022, constructed using the methodology described in Section 3. 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) Interpretation index

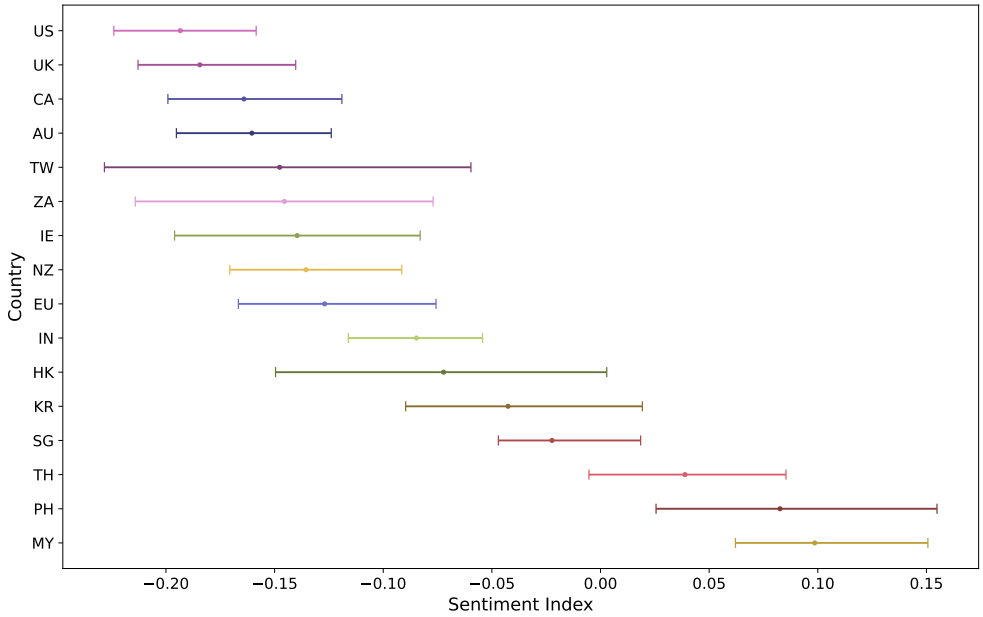
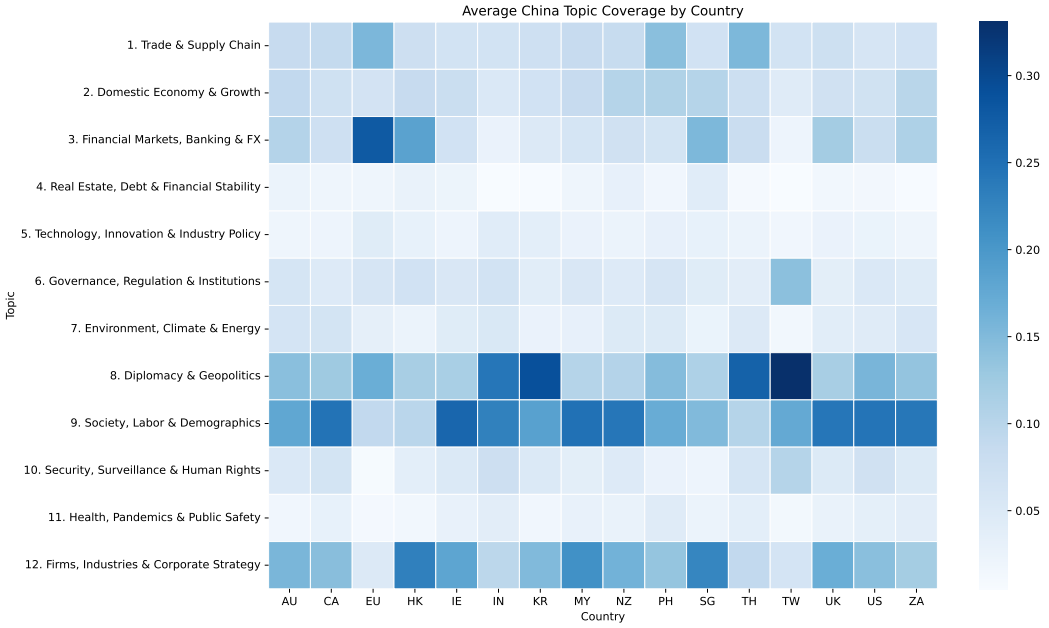


Figure 3: Cross-country 50th percentile ranges by index type

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 interpretation index, averaged over the sample period. All indices are constructed using the methodology described in Section 3.

(a) Average Topic Share Index



(b) Average Topic Interpretation Index

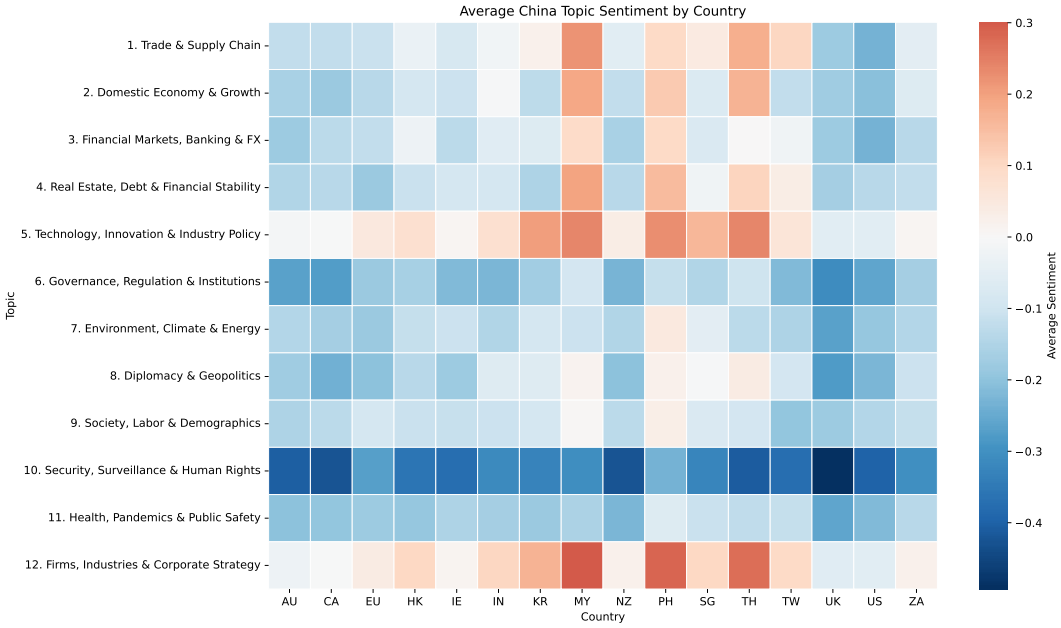


Figure 4: Average Topic Share and Sentiment by Country

Notes: This figure plots (a) the average topic share index and (b) the average topic interpretation index by country (horizontal axis) and topic (vertical axis) over the full sample period, constructed using the methodology described in Section 3. 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.

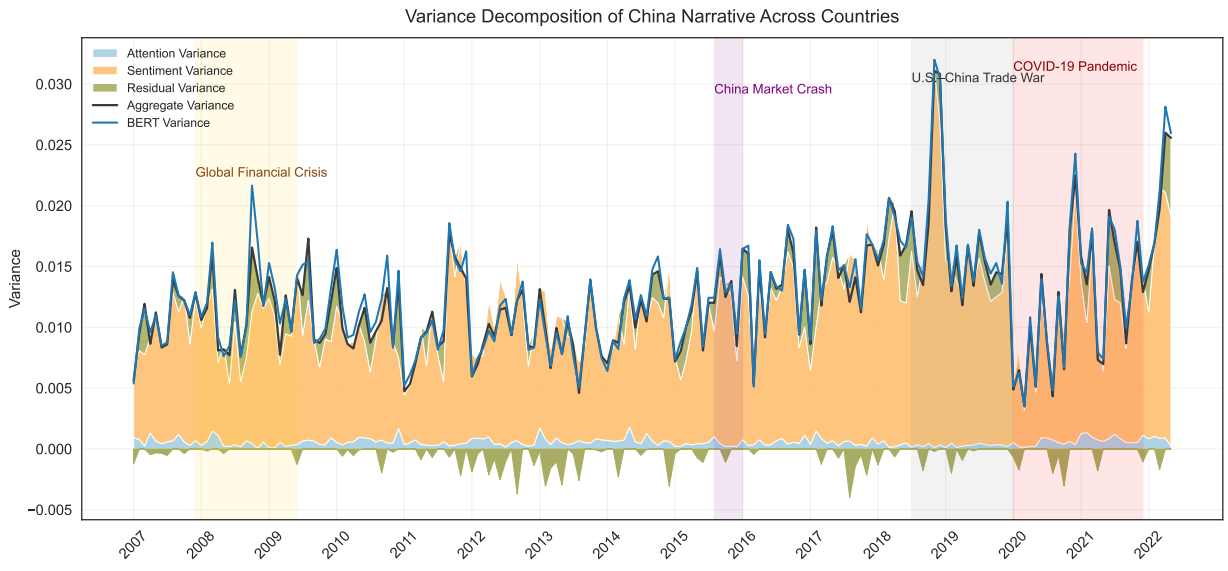
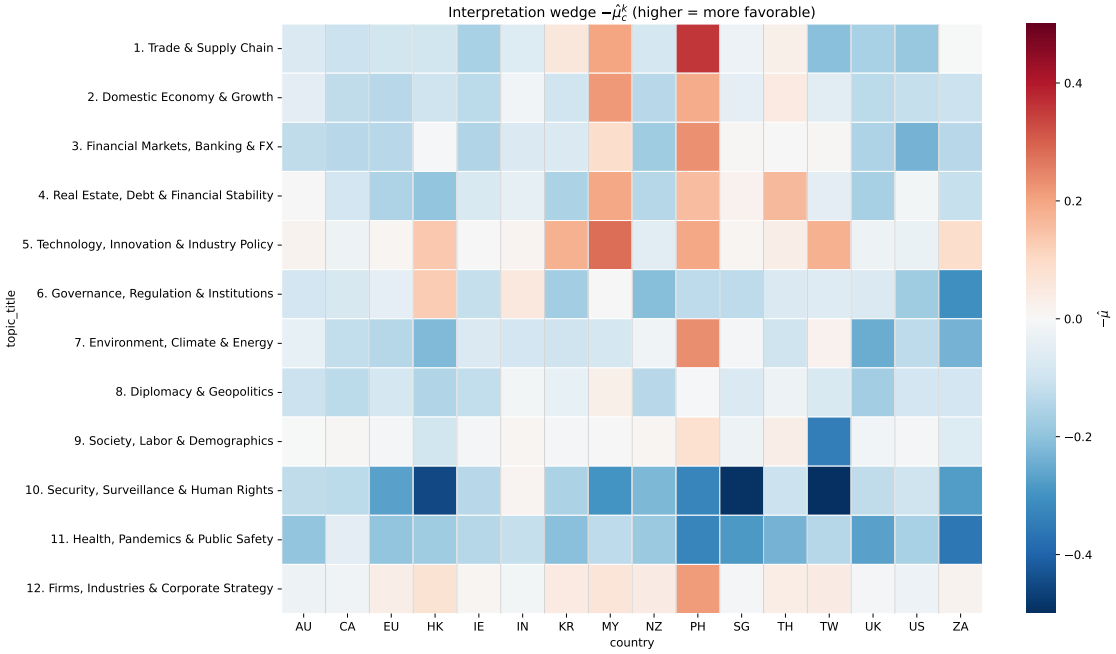


Figure 5: Variance Decomposition of China Narratives Across Countries

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 interpretation 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 4.2.

(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)

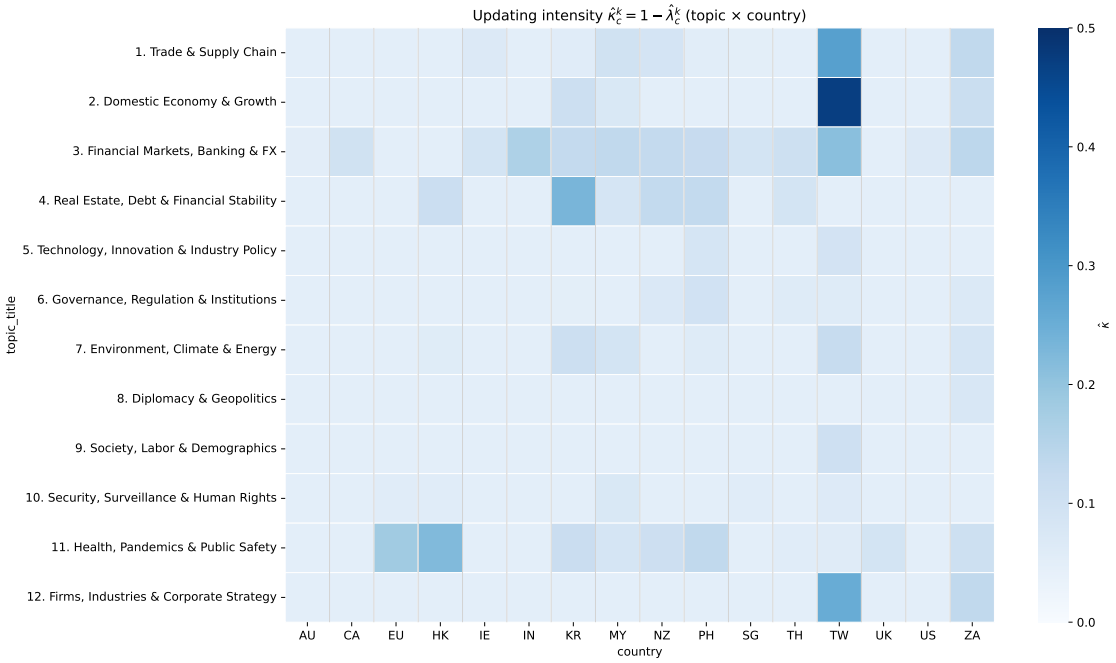


Figure 6: Country-level interpretation and updating

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.

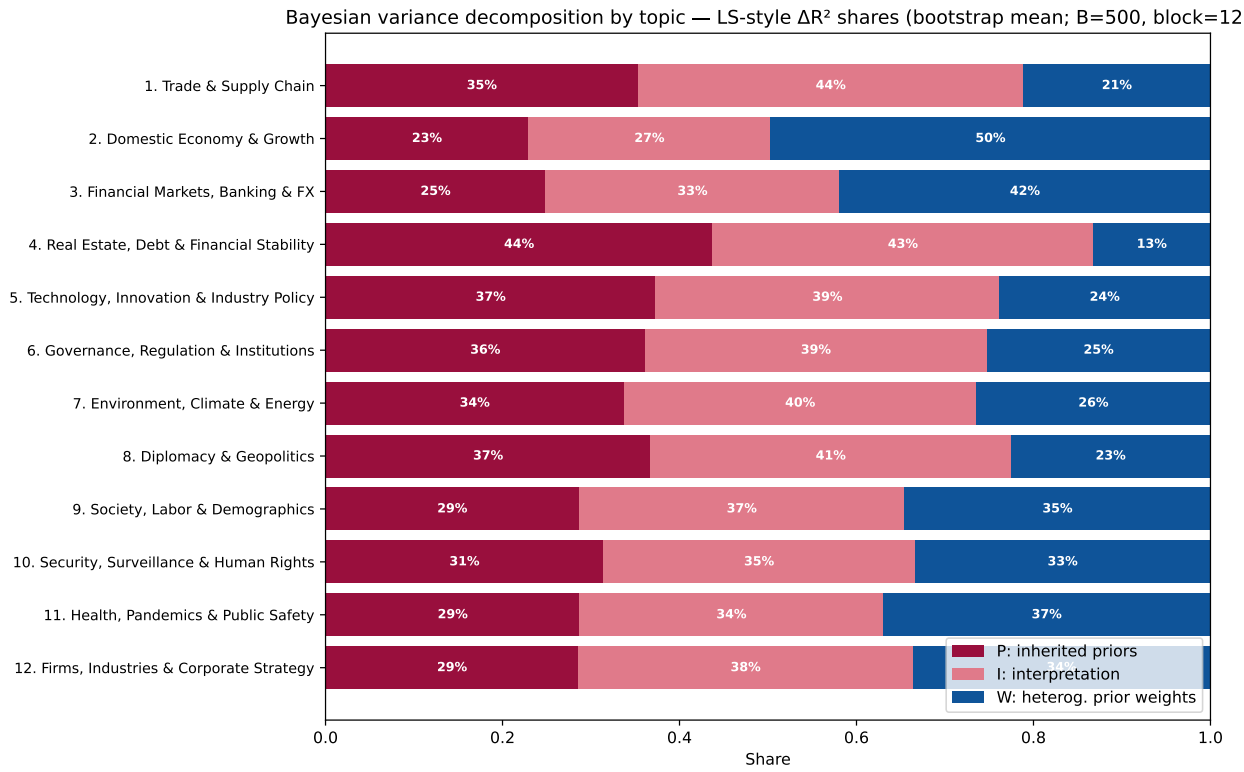


Figure 7: Belief Level vs. Updating Heterogeneity: Variance Decomposition of Topic Interpretation Index

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 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. Figure A1 further reports the results using an order-invariant Shapley decomposition method.

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 median interpretation index on China. The variable $sen^{bag}_{c,t}$ is an interpretation 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 5. 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 J.

Table 2: 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 (13) 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 interpretation 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 3: 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

Notes: This table reports estimates of equation (13) 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 5.1, and the cross language median interpretation index in Panel B as described in Section 5.1. Columns (2), (4), (6), and (8) 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), 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.

Table 4: 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 6.1. 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 (14). 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 (15). 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 5: Media narratives and capital flows using the shift-share method

Dep. Var.:	IV First Stage (monthly)	IV Second Stage			
	$sen_{c,t}$	$Flow_{ic,t}$			
	(1)	(2)	(3)	(4)	(5)
$Z_{c,t}$	0.900**				
	(0.223)				
$\widehat{sen}^{IV}_{c,t-1}$		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
Country FE	Yes				
Month FE	Yes				
Fund FE		Yes	Yes	Yes	Yes
Quarter FE		Yes	Yes	Yes	Yes
Adj. R^2	0.686	0.055	0.055	0.055	0.055
N	2486	307603	307541	307603	307541
First-stage F-stat	16.26				

Notes: This table reports estimates from the shift-share identification strategy used to construct an exogenous media narrative shock and assess its effect on capital flows, as described in Section 6.2. Column (1) reports the first-stage estimates from equation (19), which relate domestic media sentiment to the shift-share instrument constructed from non-China WSJ narrative shocks using monthly data. Columns (2)-(5) report the second-stage estimates from equation (20), which capture the effect of instrumented media sentiment on portfolio flows under alternative specifications that sequentially add macroeconomic and financial controls. Standard errors are clustered at the investor-country level in the first stage and at the fund level in the second stage and are reported in parentheses. Statistical significance is denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

Table 6: Media narratives and capital flows using the shocks from recursive VAR

Dep. Var.:	$Flow_{ic,t}$			
	(1)	(2)	(3)	(4)
$senshock_{c,t-1}$	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 (13) 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 shocks constructed using recursive vector autoregression (VAR) with Cholesky decomposition as described in Section L. Columns (2) and (4) 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) and (4) 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 7: Counterfactual analysis

Dep. Var.:	$Flow_{ic,t}$		
	(1)	(2)	(3)
$Belief_{c,t-1}$	6.792***		7.252***
	(1.507)		(1.847)
$Updating_{c,t-1}$		1.637	-0.721
		(1.122)	(1.362)
$num_{c,t-1}$	0.363	0.051	0.356
	(0.26)	(0.249)	(0.254)
Macro & Financial controls	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Adj. R^2	0.055	0.055	0.055
N	307541	307541	307541

Notes: This table reports estimates of equation (22) 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 counterfactual belief-level index and the updating index, constructed as described in Section 8. 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 8: Heterogeneous investor information and narrative-driven flows

Dep. Var.:	$Flow_{ic,t}$				
	(1)	(2)	(3)	(4)	(5)
$sen_{c,t-1}$	1.901***	1.193***	1.360***	1.157***	0.581*
	(0.284)	(0.292)	(0.267)	(0.257)	(0.268)
$sen_{c,t-1} \times EarlyStart_i$	-1.735***				
	(0.229)				
$sen_{c,t-1} \times HighShare_{i,t-1}$		-0.213			
		(0.227)			
$sen_{c,t-1} \times HighInfoStock_{i,t-1}$			-1.083***		
			(0.212)		
$sen_{c,t-1} \times GeoClose_i$				-1.904***	
				(0.471)	
$sen_{c,t-1} \times HighVol_t$					1.637**
					(0.537)
$num_{c,t-1}$	0.256	0.338	0.319	0.210	0.278
	(0.258)	(0.258)	(0.258)	(0.258)	(0.262)
$HighShare_{i,t-1}$		-10.803***			
		(0.379)			
$HIghInfoStock_{i,t-1}$			-10.785***		
			(0.498)		
Macro & Financial controls	Yes	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.055	0.059	0.057	0.055	0.055
N	307541	307541	307541	307541	307541

Notes: This table reports results from equation (21) estimated at the fund-quarter level over the period 2007Q1-2022Q1 for a sample of funds domiciled in 16 countries, with additional interaction terms. The $EarlyStart_i$ is a dummy that equals one if the investor has an initial investment quarter in Chinese assets earlier than the median of all investors in the dataset and zero if below the median. The $HighShare_{i,t}$ is a dummy that equals one if the investor has a Chinese asset share lower than the median of all investors in the dataset and zero if below the median. The $HighInfoStock_i$ is a dummy that equals one if the investor has historically accumulated substantial information about China, characterized by both higher Chinese asset holdings and an earlier investment start date than the median investors, and zero otherwise. $GeoClose_i$ is a dummy variable that equals one if the investor is domiciled in an economy geographically close to China, and zero otherwise. $HighVol_t$ is a dummy that equals one if the stock market volatility in that quarter is higher than the 75th percentile of all observed periods in our analysis. The main independent variable is interpretation index. The coefficients of $EarlyStart_i$ and $GeoClose_i$ are omitted due to collinearity with the fund fixed effects. All columns include a set of macroeconomic controls that can affect 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, the interest rate differential between China and the investor's domicile economy), and 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.

Appendix for Online Publication

Table of Contents

Part I: Theory 2

A Additional results for Private Information 2

B Additional Results for Heterogeneous Belief Levels versus Heterogeneous Updating 3
- B.1 Bayesian Learning Method Derivation . . . 3
  - B.1.1 Deriving the disagreement recursion . . . 3
  - B.1.2 Estimation: Variance decomposition . . . 4
- B.2 Counterfactual variance paths and channel attribution . . . 6
- B.3 Additional Figures for Variance Decomposition . . . 10

C Proofs 10
- C.1 Proof of Prediction 2: Expertise Attenuates Reliance on Public Interpretation 10
- C.2 Proof of Prediction 3: Destination Complexity Amplifies Reliance on Public Interpretation . . . 11

D Asymmetric Interpretation Effects and the Risk Channel 12

Part II: Data Construction 15

E Media List 15

F Interpretation Index Construction: Examples 17

G Additional Indices 18

H Cross-Language FinBERT Model Selection 19
- H.1 English . . . 19
- H.2 German . . . 19
- H.3 French . . . 20
- H.4 Spanish . . . 20

I Additional Results for Section 3.2.2 Topic-level Index 21
- I.1 Prompt Used for Topic Classification . . . 21
- I.2 Additional Figures for Topic Classification . . . 22

Part III: Empirical Results 23

J Variable Definitions 23

K The Origins of Country Belief Levels 25

L Extracting Media Narrative Shocks using VAR **30**

M Additional Results for Asymmetric and Higher-Moment Effects **31**
M.1 Asymmetric Effects of Media Narratives. 31
M.2 Higher Moment Effects of Media-Implied Risk. 32

Part I: Theory

A Additional results for Private Information

Setup. Investor i enters period t with an inherited private belief $P^k_{i,c,t-1}$ about θ^k_t, with precision $\tau^k_{i,c,t-1}$. In period t the investor acquires a new private signal:

$$x^k_{i,c,t} = \theta^k_t + \epsilon^k_{i,c,t}, \qquad \epsilon^k_{i,c,t} \sim \mathcal{N}\left(0, (r^k_{i,c,t})^{-1}\right),$$

where $r^k_{i,c,t} > 0$ is the precision of the new private signal, independent of $\epsilon^k_{c,t}$ and of the public noise.

Bayesian updating. The investor treats $P^k_{i,c,t-1}$ and $x^k_{i,c,t}$ as two independent Gaussian signals of θ^k_t. Under a diffuse or pre-integrated prior, the posterior mean is the precision-weighted average:

$$P^k_{i,c,t} = \frac{\tau^k_{i,c,t-1} P^k_{i,c,t-1} + r^k_{i,c,t} x^k_{i,c,t}}{\tau^k_{i,c,t-1} + r^k_{i,c,t}}.$$

Define the weight on the inherited belief:

$$\varphi^k_{i,c,t} \equiv \frac{\tau^k_{i,c,t-1}}{\tau^k_{i,c,t-1} + r^k_{i,c,t}} \in (0,1).$$

Then the updating equation can be written compactly as:

$$P^k_{i,c,t} = \varphi^k_{i,c,t} P^k_{i,c,t-1} + \left(1 - \varphi^k_{i,c,t}\right) x^k_{i,c,t}.$$

The updated precision is:

$$\tau^k_{i,c,t} = \tau^k_{i,c,t-1} + r^k_{i,c,t}.$$

At the boundary, when $r^k_{i,c,t} \to \infty$ (new private signal is perfectly precise), $\varphi^k_{i,c,t} \to 0$ and $P^k_{i,c,t} = x^k_{i,c,t}$: the investor discards the accumulated belief entirely. When $r^k_{i,c,t} \to 0$ (new private signal is uninformative), $\varphi^k_{i,c,t} \to 1$ and $P^k_{i,c,t} = P^k_{i,c,t-1}$: the investor relies entirely on the accumulated belief.

Combining with public interpretation. Given $P^k_{i,c,t}$ with precision $\tau^k_{i,c,t}$ and $S^k_{c,t}$ with precision $\pi^k_{c,t}$, a second round of Bayesian updating yields the portfolio belief:

$$\widehat{\theta}^k_{i,c,t} = \frac{\tau^k_{i,c,t} P^k_{i,c,t} + \pi^k_{c,t} S^k_{c,t}}{\tau^k_{i,c,t} + \pi^k_{c,t}} = \omega_{i,c,t} P^k_{i,c,t} + (1 - \omega_{i,c,t}) S^k_{c,t},$$

where $\omega_{i,c,t} = \tau^k_{i,c,t}/(\tau^k_{i,c,t} + \pi^k_{c,t})$.

Key comparative static. Since $\tau^k_{i,c,t} = \tau(\xi_i, \Omega_t)$:

$$\frac{\partial(1-\omega_{i,c,t})}{\partial \tau^k_{i,c,t}} = -\frac{\pi^k_{c,t}}{(\tau^k_{i,c,t}+\pi^k_{c,t})^2} < 0.$$

Higher accumulated private precision unambiguously reduces the weight on source-country public interpretation, directly motivating Prediction 2.

B Additional Results for Heterogeneous Belief Levels versus Heterogeneous Updating

B.1 Bayesian Learning Method Derivation

B.1.1 Deriving the disagreement recursion

Sentiment revision. Define the sentiment revision:

$$\Delta S^k_{c,t} \equiv S^k_{c,t} - S^k_{c,t-1} = (1-\lambda^k_{c,t})\Big(L^k_t - \mu^k_{c,t} - S^k_{c,t-1}\Big). \tag{A1}$$

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

$$\bar{S}^k_t \equiv \mathbb{E}_c[S^k_{c,t}], \quad \sigma^2_{S|t,k} \equiv \mathrm{Var}_c(S^k_{c,t}), \quad \lambda^k_t \equiv \mathbb{E}_c[\lambda^k_{c,t}], \quad \sigma^2_{\lambda|t,k} \equiv \mathrm{Var}_c(\lambda^k_{c,t}), \tag{A2}$$

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 (A1) yields the mean revision:

$$\Delta \bar{S}^k_t \equiv \bar{S}^k_t - \bar{S}^k_{t-1} = (1-\lambda^k_t)\Big(L^k_t - \mu^k_t - \bar{S}^k_{t-1}\Big). \tag{A3}$$

Equation (A3) 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:

$$\mathrm{Cov}_c(S^k_{c,t-1}, \lambda^k_{c,t}) = \mathrm{Cov}_c(S^k_{c,t-1}, \mu^k_{c,t}) = \mathrm{Cov}_c(\lambda^k_{c,t}, \mu^k_{c,t}) = 0. \tag{A4}$$

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

Variance algebra From (1),

$$\sigma^2_{S|t,k} = \mathrm{Var}_c\Big(\lambda^k_{c,t} S^k_{c,t-1}\Big) + \mathrm{Var}_c\Big((1-\lambda^k_{c,t})(L^k_t - \mu^k_{c,t})\Big) + 2\,\mathrm{Cov}_c\Big(\lambda^k_{c,t} S^k_{c,t-1},\ (1-\lambda^k_{c,t})(L^k_t - \mu^k_{c,t})\Big). \tag{A5}$$

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, (A4) implies:

$$\mathrm{Var}_c(\lambda^k_{c,t} S^k_{c,t-1}) = \sigma^2_{\lambda|t,k}\sigma^2_{S|t-1,k} + \sigma^2_{\lambda|t,k}(\bar{S}^k_{t-1})^2 + \sigma^2_{S|t-1,k}(\lambda^k_t)^2, \tag{A6}$$

$$\mathrm{Var}_c\Big((1-\lambda^k_{c,t})(L^k_t - \mu^k_{c,t})\Big) = \sigma^2_{\lambda|t,k}\sigma^2_{\mu|t,k} + \sigma^2_{\lambda|t,k}(L^k_t - \mu^k_t)^2 + \sigma^2_{\mu|t,k}(1-\lambda^k_t)^2, \tag{A7}$$

$$\mathrm{Cov}_c\Big(\lambda^k_{c,t} S^k_{c,t-1},\ (1-\lambda^k_{c,t})(L^k_t - \mu^k_{c,t})\Big) = -(L^k_t - \mu^k_t)\,\bar{S}^k_{t-1}\,\sigma^2_{\lambda|t,k}. \tag{A8}$$

Substituting (A6)–(A8) into (A5) yields:

$$\sigma^2_{S|t,k} = \sigma^2_{S|t-1,k}\Big(\sigma^2_{\lambda|t,k} + (\lambda^k_t)^2\Big) + \sigma^2_{\mu|t,k}\Big(\sigma^2_{\lambda|t,k} + (1-\lambda^k_t)^2\Big) + \sigma^2_{\lambda|t,k}\Big(L^k_t - \mu^k_t - \bar{S}^k_{t-1}\Big)^2. \tag{A9}$$

Disagreement recursion. From (A3), $(L^k_t - \mu^k_t - \bar{S}^k_{t-1}) = \Delta\bar{S}^k_t/(1-\lambda^k_t)$. Substituting into (A9) gives the LS disagreement recursion:

$$\sigma^2_{S|t,k} = \sigma^2_{S|t-1,k}\Big(\sigma^2_{\lambda|t,k} + (\lambda^k_t)^2\Big) + \sigma^2_{\mu|t,k}\Big(\sigma^2_{\lambda|t,k} + (1-\lambda^k_t)^2\Big) + \sigma^2_{\lambda|t,k}\left[\frac{\Delta\bar{S}^k_t}{1-\lambda^k_t}\right]^2. \tag{A10}$$

Equation (A10) 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 ($\sigma^2_{\mu|t,k}$).

B.1.2 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 (A10). We nonetheless interpret L^k_t as our constructed global topic information index.

Estimation of λ via the forecast-revision regression.

Random-coefficient restriction and revision regression. Rewriting (1) as

$$S^k_{c,t} = \lambda^k_{c,t} S^k_{c,t-1} + \varepsilon^k_{c,t}, \qquad \varepsilon^k_{c,t} \equiv (1-\lambda^k_{c,t})(L^k_t - \mu^k_{c,t}), \tag{A11}$$

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

$$\lambda_{c,t}^k = \lambda^k + v_c^k, \qquad \mathbb{E}_c[v_c^k] = 0, \qquad \mathrm{Var}_c(v_c^k) = \sigma_{\lambda|k}^2, \tag{A12}$$

with v_c^k mutually independent across c and independent of $(S_{c,t-1}^k, \mu_{c,t}^k)$ in the cross section. As in LS, we use the revision regression to avoid spurious dynamics:

$$\Delta S_{c,t}^k = \beta^k S_{c,t-1}^k + u_{c,t}^k, \qquad \beta^k \equiv \lambda^k - 1, \tag{A13}$$

where $u_{c,t}^k = \varepsilon_{c,t}^k + v_c^k S_{c,t-1}^k$. Under (A4) and (A12), $u_{c,t}^k$ is uncorrelated with $S_{c,t-1}^k$.

Common correlated effects (CCE) augmentation. To allow for cross-country correlation from common aggregate shocks and topic-wide narratives, we follow LS and augment (A13) with Pesaran (2006) CCE controls:

$$\Delta S_{c,t}^k = \beta^k S_{c,t-1}^k + \boldsymbol{c}^{k\prime} \boldsymbol{z}_t^k + \eta_{c,t}^k, \qquad \boldsymbol{z}_t^k \equiv \left(1,\ \bar{S}_{t-1}^k,\ \Delta \bar{S}_t^k\right). \tag{A14}$$

We estimate (A14) separately by topic k using pooled OLS across all (c,t) observations available. We then recover

$$\widehat{\lambda}^k \equiv 1 + \widehat{\beta}^k, \qquad \widehat{\kappa}^k \equiv 1 - \widehat{\lambda}^k \quad \text{(weight on new public information)}. \tag{A15}$$

For numerical stability in the variance-step below, and following LS, we impose an upper bound $\widehat{\lambda}^k \le 0.95$.[31]

Estimation of σ_λ^2 and σ_μ^2.

Regression form of the variance recursion. Plugging $\widehat{\lambda}^k$ into (A10) and rearranging yields the LS estimating equation (the analogue of LS Eq. (21) with calendar time in place of forecast horizon):

$$\sigma_{S|t,k}^2 - (\widehat{\lambda}^k)^2 \sigma_{S|t-1,k}^2 = \sigma_{\lambda|k}^2 \sigma_{\mu|k}^2 + \sigma_{\lambda|k}^2 \left[\sigma_{S|t-1,k}^2 + \left(\frac{\Delta \bar{S}_t^k}{1 - \widehat{\lambda}^k} \right)^2 \right] + \sigma_{\mu|k}^2 (1 - \widehat{\lambda}^k)^2 + e_{t,k}. \tag{A16}$$

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

Define:

$$Y_{t,k} \equiv \sigma_{S|t,k}^2 - (\widehat{\lambda}^k)^2 \sigma_{S|t-1,k}^2, \qquad X_{t,k} \equiv \sigma_{S|t-1,k}^2 + \left(\frac{\Delta \bar{S}_t^k}{1 - \widehat{\lambda}^k} \right)^2. \tag{A17}$$

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

Under the time-invariance restriction, the two terms involving $\sigma^2_{\mu|k}$ in (A16) can be collected into a single intercept:

$$Y_{t,k} = \sigma^2_{\lambda|k} X_{t,k} + \alpha_k + e_{t,k}, \qquad \alpha_k \equiv \sigma^2_{\mu|k}\Big(\sigma^2_{\lambda|k} + (1-\widehat{\lambda}^k)^2\Big). \tag{A18}$$

Equation (A18) is *linear* in the estimable regression coefficients $\big(\sigma^2_{\lambda|k}, \alpha_k\big)$; $\sigma^2_{\mu|k}$ is then recovered algebraically from α_k.

OLS estimation. We estimate (A18) *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).[32] 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 (A18),

$$\widehat{\sigma}^2_{\mu|k} \equiv \frac{\widehat{\alpha}_k}{\widehat{\sigma}^2_{\lambda|k} + (1-\widehat{\lambda}^k)^2}. \tag{A19}$$

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).

B.2 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 (A10) forward in calendar time:

$$\widehat{\sigma}^2_{S|t,k} = \widehat{\sigma}^2_{S|t-1,k}\Big(\widehat{\sigma}^2_{\lambda|k} + (\widehat{\lambda}^k)^2\Big) + \widehat{\sigma}^2_{\mu|k}\Big(\widehat{\sigma}^2_{\lambda|k} + (1-\widehat{\lambda}^k)^2\Big) + \widehat{\sigma}^2_{\lambda|k}\left[\frac{\Delta \bar{S}^k_t}{1-\widehat{\lambda}^k}\right]^2, \tag{A20}$$

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 $\sigma^2_{\mu|k}$.

[32] 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*, (A18) 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 (A10) 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 $t > t_0$,

$$v_{t,k}(A) = v_{t-1,k}(A)\Big(\mathbb{1}\{W \in A\}\widehat{\sigma}^2_{\lambda|k} + (\widehat{\lambda}^k)^2\Big) + \mathbb{1}\{I \in A\}\widehat{\sigma}^2_{\mu|k}\Big(\mathbb{1}\{W \in A\}\widehat{\sigma}^2_{\lambda|k} + (1-\widehat{\lambda}^k)^2\Big) + \mathbb{1}\{W \in A\}\widehat{\sigma}^2_{\lambda|k}\left[\frac{\Delta\bar{S}^k_t}{1-\widehat{\lambda}^k}\right]^2. \tag{A21}$$

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 (A21)).

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 $t = t_0, \ldots, T_k$:

$$D_{t,k} \;=\; a^k_A \;+\; b^k_A\, v_{t,k}(A) \;+\; u_{t,k}(A), \tag{A22}$$

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

$$\Delta R^2_{k,j} \;\equiv\; R^2_{k,\text{full}} - R^2_{k,-j}, \qquad j \in \{P, W, I\}. \tag{A23}$$

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

$$\text{Share}_{k,j}^{\Delta R^2} \equiv \frac{\Delta R_{k,j}^2}{\Delta R_{k,P}^2 + \Delta R_{k,W}^2 + \Delta R_{k,I}^2}, \qquad j \in \{P, W, I\}. \tag{A24}$$

Because $v_{t,k}(A)$ is generated by a nonlinear recursion and (A22) is only a fit diagnostic, $\Delta R_{k,j}^2$ can occasionally be small or negative in finite samples. In that case, we follow the standard convention of setting negative $\Delta R_{k,j}^2$ to zero prior to normalization and interpret (A24) 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 $\text{Share}_{k,j}^{\Delta R^2}$ 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

$$V_k(A) \equiv R_k^2(A),$$

where $R_k^2(A)$ is obtained from (A22). (With an intercept in (A22), $R_k^2(\emptyset) = 0$ 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

$$\phi_k(j) = \sum_{A \subseteq \mathcal{N} \setminus \{j\}} \frac{|A|!\,(m - |A| - 1)!}{m!} \Big(V_k(A \cup \{j\}) - V_k(A) \Big). \tag{A25}$$

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 $m = 3$, (A25) reduces to the following

convenient closed forms (useful for coding):

$$\phi_k(P) = \frac{1}{3}\Big(R^2_k(P) - R^2_k(\emptyset)\Big) + \frac{1}{6}\Big(R^2_k(P,W) - R^2_k(W)\Big) + \frac{1}{6}\Big(R^2_k(P,I) - R^2_k(I)\Big) + \frac{1}{3}\Big(R^2_k(P,W,I) - R^2_k(W,I)\Big), \tag{A26}$$

$$\phi_k(W) = \frac{1}{3}\Big(R^2_k(W) - R^2_k(\emptyset)\Big) + \frac{1}{6}\Big(R^2_k(P,W) - R^2_k(P)\Big) + \frac{1}{6}\Big(R^2_k(W,I) - R^2_k(I)\Big) + \frac{1}{3}\Big(R^2_k(P,W,I) - R^2_k(P,I)\Big), \tag{A27}$$

$$\phi_k(I) = \frac{1}{3}\Big(R^2_k(I) - R^2_k(\emptyset)\Big) + \frac{1}{6}\Big(R^2_k(P,I) - R^2_k(P)\Big) + \frac{1}{6}\Big(R^2_k(W,I) - R^2_k(W)\Big) + \frac{1}{3}\Big(R^2_k(P,W,I) - R^2_k(P,W)\Big). \tag{A28}$$

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

$$\text{Share}^{\text{Shap}}_{k,j} \equiv \frac{\phi_k(j)}{\phi_k(P) + \phi_k(W) + \phi_k(I)} = \frac{\phi_k(j)}{R^2_{k,\text{full}}}, \qquad j \in \{P, W, I\}. \tag{A29}$$

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 (A29) as an approximate decomposition. We also apply the same moving-block bootstrap over months.

B.3 Additional Figures for Variance Decomposition

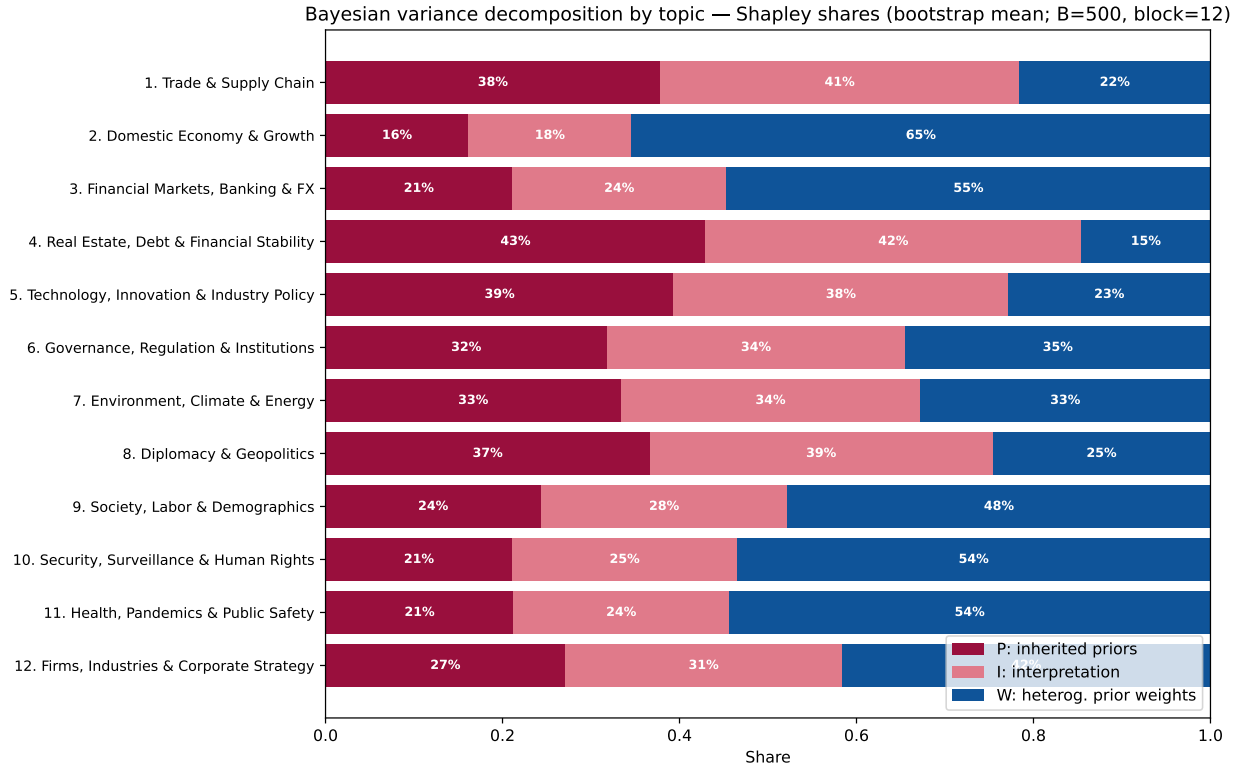


Figure A1: Belief Level vs. Updating Heterogeneity: Variance Decomposition of Topic Interpretation Index (Shapley Decomposition Method)

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 results use an order-invariant Shapley decomposition, which allocates the model-implied variance surplus across components by averaging marginal contributions over all possible inclusion orders.

C Proofs

C.1 Proof of Prediction 2: Expertise Attenuates Reliance on Public Interpretation

Using $h_k \equiv \pi^k_{c,t} + \tau^k_{i,c,t}$ and $D \equiv \sigma^2_{\varepsilon,d,t} + \sum_j \beta_j^2/h_j$, note that D does not depend on $S^k_{c,t}$, so differentiating equation (8) with respect to $S^k_{c,t}$ yields:

$$\frac{\partial a^{d,*}_{i,c,t}}{\partial S^k_{c,t}} = \frac{\beta_k(1-\omega_{i,c,t})}{\gamma_i D}.$$

Substituting $(1-\omega_{i,c,t}) = \pi^k_{c,t}/h_k$:

$$\frac{\partial a^{d,*}_{i,c,t}}{\partial S^k_{c,t}} = \frac{\beta_k \pi^k_{c,t}}{\gamma_i h_k D}.$$

Differentiating with respect to $\tau^k_{i,c,t}$ via the chain rule, since $\partial h_k/\partial \tau^k_{i,c,t} = 1$ and $\partial D/\partial \tau^k_{i,c,t} = -\beta_k^2/h_k^2$:

$$\frac{\partial a^{d,*}_{i,c,t}}{\partial S^k_{c,t}\partial \tau^k_{i,c,t}} = \frac{\partial}{\partial \tau^k_{i,c,t}}\left[\frac{\beta_k \pi^k}{\gamma_i h_k D}\right] = -\frac{\beta_k \pi^k}{\gamma_i (h_k D)^2}\left(D - \frac{\beta_k^2}{h_k}\right) = -\frac{\beta_k \pi^k}{\gamma_i (h_k D)^2}\left(\sigma^2_{\varepsilon,d,t} + \sum_{j\neq k}\frac{\beta_j^2}{h_j}\right).$$

Following that and since $\frac{\partial \tau^k_{i,c,t}}{\partial \xi_i} > 0$:

$$\frac{\partial^2 a^{d,*}_{i,c,t}}{\partial S^k_{c,t}\partial \xi_i} = \frac{\partial a^{d,*}_{i,c,t}}{\partial S^k_{c,t}\partial \tau^k_{i,c,t}}\frac{\partial \tau^k_{i,c,t}}{\partial \xi_i} = -\frac{\beta_k \pi^k}{\gamma_i (h_k D)^2}\left(\sigma^2_{\varepsilon,d,t} + \sum_{j\neq k}\frac{\beta_j^2}{h_j}\right)\frac{\partial \tau^k_{i,c,t}}{\partial \xi_i}.$$

The overall expression is negative if and only if:

$$\sigma^2_{\varepsilon,d,t} + \sum_{j\neq k}\frac{\beta_j^2}{h_j} > 0,$$

which holds whenever there is residual payoff risk ($\sigma^2_{\varepsilon,d,t} > 0$) or more than one payoff-relevant topic. This condition fails only in the knife-edge case of a single topic ($K = 1$) with no residual return risk, which has no empirical content. □

C.2 Proof of Prediction 3: Destination Complexity Amplifies Reliance on Public Interpretation

Using the same shorthand $h_k \equiv \pi^k_{c,t} + \tau^k_{i,c,t}$ and $D \equiv \sigma^2_{\varepsilon,d,t} + \sum_j \beta_j^2/h_j$ as in Appendix C.1, and noting that D does not depend on $S^k_{c,t}$, differentiating equation (8) with respect to $S^k_{c,t}$ and substituting $(1-\omega_{i,c,t}) = \pi^k_{c,t}/h_k$ gives:

$$\frac{\partial a^{d,*}_{i,c,t}}{\partial S^k_{c,t}} = \frac{\beta_k \pi^k_{c,t}}{\gamma_i h_k D}.$$

Differentiating with respect to Ω_t conditional on $\pi^k_{c,t}$ and since $\frac{\partial \tau^k_{i,c,t}}{\partial \Omega_i} < 0$:

$$\left.\frac{\partial^2 a^{d,*}_{i,c,t}}{\partial S^k_{c,t}\partial \Omega_t}\right|_{\pi^k_{c,t}} = \frac{\partial a^{d,*}_{i,c,t}}{\partial S^k_{c,t}\partial \tau^k_{i,c,t}}\frac{\partial \tau^k_{i,c,t}}{\partial \Omega_t} = -\frac{\beta_k \pi^k}{\gamma_i (h_k D)^2}\left(\sigma^2_{\varepsilon,d,t} + \sum_{j\neq k}\frac{\beta_j^2}{h_j}\right)\frac{\partial \tau^k_{i,c,t}}{\partial \Omega_t}.$$

The overall expression is positive if and only if:

$$\sigma^2_{\varepsilon,d,t} + \sum_{j \neq k} \frac{\beta_j^2}{h_j} > 0,$$

which holds under the same near-universal regularity condition as Prediction 2: residual payoff risk or more than one payoff-relevant topic. Higher destination complexity therefore unambiguously raises allocation sensitivity to source-country public interpretation under this condition. □

D Asymmetric Interpretation Effects and the Risk Channel

This section derives the theoretical results. Empirical tests of these predictions are reported in Appendix M.

Augmented perceived return variance. The baseline posterior return variance derived in Section 2.4 captures residual return risk and fundamental uncertainty:

$$\widehat{\sigma}^{2,d}_{i,c,t} = \sigma^2_{\varepsilon,d,t} + \sum_k \beta_k^2 v^k_{i,c,t}, \qquad v^k_{i,c,t} = \frac{1}{\pi^k_{c,t} + \tau^k_{i,c,t}}.$$

Source-country public interpretation conveys not only information about destination fundamentals through $S^k_{c,t}$ but also signals about destination uncertainty through its negative framing and risk content. Consistent with the negativity bias and ambiguity aversion documented in Soroka (2006) and Epstein and Schneider (2008), we augment the perceived return variance as follows:

$$\widetilde{\sigma}^{2,d}_{i,c,t} = \widehat{\sigma}^{2,d}_{i,c,t} + \phi_- N_{c,t} + \phi_R R_{c,t}, \qquad \phi_-, \phi_R > 0, \tag{A30}$$

where $N_{c,t} \geq 0$ denotes the adverse interpretation intensity of source-country public interpretation and $R_{c,t} \geq 0$ denotes its risk content. Higher adverse interpretation or heightened risk framing raises perceived return variance independently of average sentiment $S^k_{c,t}$. Note that $N_{c,t}$ and $S^k_{c,t}$ are related but distinct: $S^k_{c,t}$ captures the average tone of public interpretation while $N_{c,t}$ captures specifically the intensity of adverse framing, consistent with the separate positive and negative indices constructed empirically in Section 3.

Augmented portfolio demand. Substituting (A30) into the mean-variance demand equation (2.4), the investor's augmented portfolio demand is:

$$\widetilde{a}^{d,*}_{i,c,t} = \frac{\widehat{\mu}^d_{i,c,t}}{\gamma_i \widetilde{\sigma}^{2,d}_{i,c,t}} = \frac{\alpha_d + \sum_k \beta_k \left(\omega_{i,c,t} P^k_{i,c,t} + (1 - \omega_{i,c,t}) S^k_{c,t}\right)}{\gamma_i \left(\widehat{\sigma}^{2,d}_{i,c,t} + \phi_- N_{c,t} + \phi_R R_{c,t}\right)}. \tag{A31}$$

We now derive the two predictions that follow from this augmented demand equation.

Prediction 4 (Asymmetric Effects of Interpretations). *Let $S_{c,t}^{k,+}$ and $S_{c,t}^{k,-}$ denote the positive and negative components of $S_{c,t}^{k}$ respectively, so that $S_{c,t}^{k} = S_{c,t}^{k,+} - S_{c,t}^{k,-}$ with $S_{c,t}^{k,+}, S_{c,t}^{k,-} \geq 0$, and let $N_{c,t}$ denote adverse interpretation intensity which is increasing in $S_{c,t}^{k,-}$. The allocation response to a positive interpretation shift operates only through the expected return channel:*

$$\frac{\partial \widetilde{a}_{i,c,t}^{d,*}}{\partial S_{c,t}^{k,+}} = \frac{\beta_k(1-\omega_{i,c,t})}{\gamma_i \widetilde{\sigma}_{i,c,t}^{2,d}} > 0.$$

The allocation response to a negative interpretation shift operates through both the expected return channel and the variance channel simultaneously:

$$\frac{\partial \widetilde{a}_{i,c,t}^{d,*}}{\partial S_{c,t}^{k,-}} = -\frac{\beta_k(1-\omega_{i,c,t})}{\gamma_i \widetilde{\sigma}_{i,c,t}^{2,d}} - \frac{\widehat{\mu}_{i,c,t}^{d}}{\gamma_i (\widetilde{\sigma}_{i,c,t}^{2,d})^2} \cdot \phi_{-} \frac{\partial N_{c,t}}{\partial S_{c,t}^{k,-}} < 0,$$

where the first term is the expected return channel and the second term is the variance channel, both of which reduce allocation. Since the variance channel is absent for positive shifts but present for negative shifts, adverse interpretations have larger allocation effects than favorable interpretations of equal magnitude:

$$\left| \frac{\partial \widetilde{a}_{i,c,t}^{d,*}}{\partial S_{c,t}^{k,-}} \right| > \left| \frac{\partial \widetilde{a}_{i,c,t}^{d,*}}{\partial S_{c,t}^{k,+}} \right|.$$

We do not impose a symmetric response to positive interpretation: favorable sentiment raises expected returns through $S_{c,t}^{k}$ without a corresponding reduction in $\widetilde{\sigma}_{i,c,t}^{2,d}$, consistent with Soroka (2006) and Epstein and Schneider (2008).

Prediction 5 (Source-Country Interpretation of Risk Reduces Allocation). *Source-country public interpretation conveys not only information about destination fundamentals θ_t^k but also signals about destination-country uncertainty captured by $R_{c,t}$. From equation (A30), heightened risk content raises $\widetilde{\sigma}_{i,c,t}^{2,d}$ independently of average sentiment $S_{c,t}^{k}$:*

$$\frac{\partial \widetilde{\sigma}_{i,c,t}^{2,d}}{\partial R_{c,t}} = \phi_R > 0.$$

For an investor holding a positive expected destination return ($\widehat{\mu}_{i,c,t}^{d} > 0$, corresponding to an interior long position in the destination asset), the allocation response to heightened risk content is:

$$\frac{\partial \widetilde{a}_{i,c,t}^{d,*}}{\partial R_{c,t}} = -\frac{\widehat{\mu}_{i,c,t}^{d}}{\gamma_i (\widetilde{\sigma}_{i,c,t}^{2,d})^2} \cdot \phi_R < 0.$$

Heightened risk content in source-country public interpretation therefore reduces desired destination exposure, operating as a second-moment channel distinct from the expected-return

channel in Predictions 1–4.[33] *The magnitude of this effect is increasing in* ϕ_R *and decreasing in* $\widetilde{\sigma}^{2,d}_{i,c,t}$*: investors who already perceive high return variance are less sensitive to additional risk signals at the margin.*

[33]The condition $\widehat{\mu}^{d}_{i,c,t} > 0$ holds for funds with a positive expected excess return on Chinese assets, which is the relevant case for the open-ended mutual funds in our sample that hold long positions in Chinese assets. Short positions ($\widehat{\mu}^{d}_{i,c,t} < 0$) would reverse the sign, but are not common under the investment mandates of the funds in our dataset.

Part II: Data Construction

E 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.

F Interpretation Index 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.

G 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 Interpretation Index. We follow the similar methodology of Flynn and Sastry (2024) and Hassan, Hollander, Van Lent and Tahoun (2019) to construct the interpretation 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 interpretation 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.

H 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.

H.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.

H.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.

H.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.

H.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. They provide a unified multilingual FinBERT framework for cross-country sentiment analysis in financial research.

I Additional Results for Section 3.2.2 Topic-level Index

I.1 Prompt Used for Topic 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.

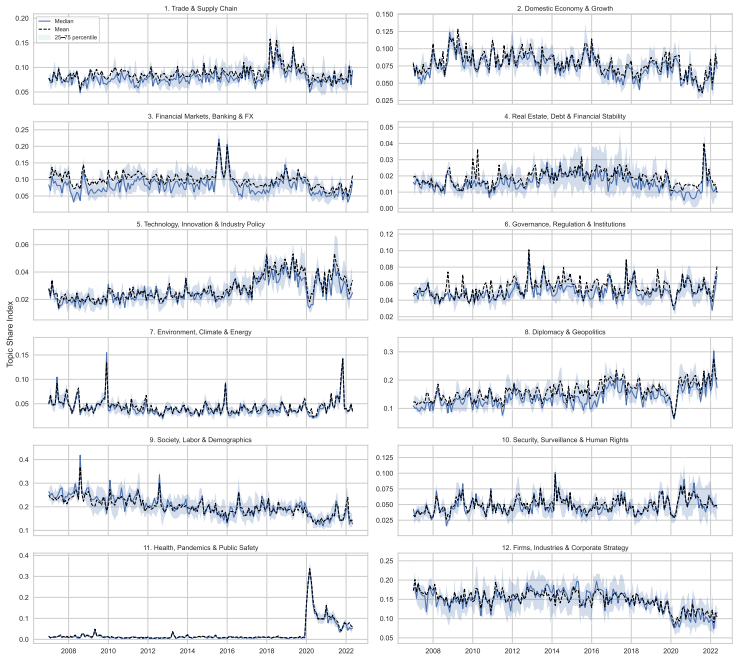
11. Health, Pandemics \& Public Safety
   Disease outbreaks, vaccines, COVID-19, medical systems, public safety.

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>"}
```

I.2 Additional Figures for Topic Classification

(a) Topic share index



(b) Topic interpretation index

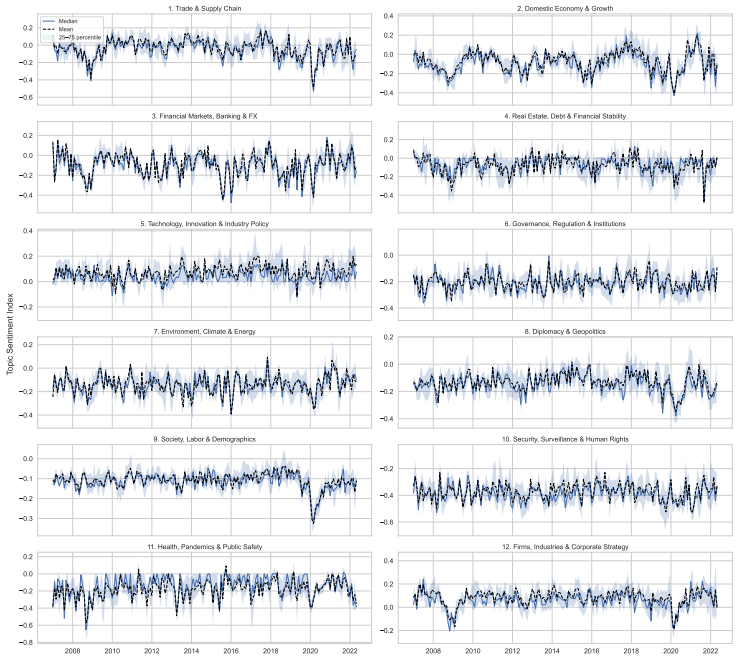


Figure A2: Topic Share and Sentiment Indices

Notes: This figure plots the cross-country distribution of (a) the topic share index and (b) the topic interpretation index for China over the period 2007–2022, constructed using the methodology described in Section 3. 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).

Part II: Data Construction

J Variable Definitions

Table A4 shows the definitions and source of variables used in equation (13).

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 (13).

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 relationship 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 K.

K The Origins of Country Belief Levels

In this section, we 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 4.3.1. We relate these belief levels to slow-moving (or time-invariant) country characteristics that proxy for economic relationship, geopolit-

ical alignment, proximity, institutional environment, and market depth and financial similarity.

Our specification pools all topics and includes topic fixed effects:

$$-\widehat{\mu}_c^k = \beta' X_c + \delta_k + \varepsilon_{c,k}, \tag{A32}$$

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).[34] 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 relationship 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)

[34] 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.

equity turnover ($Turnover_c$), and (iii) return-based similarity to China's stock market ($RetCorrCN_c$).

Results. Table A7 reports the cross-country determinants of narrative belief levels about China. Economic relationship 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.

Table A7: The origins of national belief levels

Dep. Var.:	$-\widehat{\mu}_c^k$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Economic relationship with China:								
$TradeCN_c$	0.034							0.449***
	(0.057)							(0.028)
$TradeShare_c$	-0.381							0.547***
	(0.226)							(0.095)
Geopolitical alignment:								
$AgreeCN_c$		0.207**						
		(0.064)						
BRI_c		-0.017						-0.112***
		(0.024)						(0.013)
Proximity:								
$DistCN_c$			-0.023					-0.018**
			(0.023)					(0.004)
$ContCN_c$			0.000					-0.085
			(0.042)					(0.047)
$ComLangCN_c$			-0.025					
			(0.061)					
Institutional quality:								
$RuleLaw_c$				-0.049*				-0.058**
				(0.019)				(0.014)
$GovEff_c$					-0.045			
					(0.023)			
$RegQuality_c$						-0.046*		
						(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)
$RetCorrCN_c$							0.327	0.387
							(0.301)	(0.177)
Constant	-0.059*	-0.183**	0.124	-0.016	-0.016	-0.015	-0.135	0.125***
	(0.022)	(0.043)	(0.199)	(0.028)	(0.035)	(0.029)	(0.095)	0.000
Topic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.334	0.421	0.310	0.424	0.382	0.406	0.350	0.520
N	180	168	192	180	180	180	168	168

Notes: This table reports estimates of equation (A32). The dependent variable is the estimated country-specific narrative belief level across 12 topics. The independent variables capture five conceptual dimensions: economic relationship 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 A8 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 A8: 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 (A32).

L Extracting Media Narrative Shocks using VAR

For a third approach to identify causal effects, we isolate media narrative shocks that are orthogonal to macroeconomic and financial fundamentals. To isolate the narrative component ($narrativeshocks_{c,t}$) from the interpretation index, we estimate the following equation:

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

where $sen_{c,t}$ is the interpretation 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 decomposition to identify $\varepsilon_{c,t}$. The baseline VAR specification is:

$$\boldsymbol{A}\boldsymbol{Y_t} = \sum_{j=1}^{p} \boldsymbol{C_j}\boldsymbol{Y_{t-j}} + \boldsymbol{\varepsilon_t}$$

where $\boldsymbol{A}$ and $\boldsymbol{C_j}$ 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 $s \neq t$. 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 $\boldsymbol{Y_t}$ includes the following variables, all in log form unless otherwise noted: the constructed median interpretation 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.[35]

We order the variables in $\boldsymbol{Y_t}$ as follows based on standard macro-financial assumptions: (1) the media index, under the assumption that narrative shocks can contemporaneously

[35]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 interpretation index is centered around zero; therefore, taking logs is neither meaningful nor feasible.

affect all other variables, but are not themselves contemporaneously affected by them; (2) stock returns, as they react immediately to news and are considered highly responsive; (3) the volatility index, as it typically reacts to returns but exhibits more persistence and slower transmission to other variables; (4) the exchange rate index, reflecting its sensitivity to financial variables while assuming it does not contemporaneously influence real economic activities; (5) the CPI, based on its slower-moving nature and limited immediate responsiveness to financial shocks; (6) the industrial production index, as it is assumed to be the most sluggish among the variables and does not respond contemporaneously to any of the preceding shocks.[36]

The first element of ε_t, corresponding to the media index, is interpreted as the media narrative shock. We estimate this 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. Table 6 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.

M Additional Results for Asymmetric and Higher-Moment Effects

Theory for these predictions is derived in Appendix D.s

M.1 Asymmetric Effects of Media Narratives.

Prediction 4 implies $|\partial a/\partial S^{k,-}| > |\partial a/\partial S^{k,+}|$. Motivated by this prediction, we decompose the interpretation index into positive and negative components.

We construct two separate indices based on word frequencies.[37] Using the positive and negative word lists from Loughran and McDonald (2011), the positive media narrative

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

[37] 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.

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 country. The negative media narrative index, $neg_{c,t}$, is constructed analogously using negative word frequencies.[38]

Panel A of Table A9 reports the results from specifications that jointly include the positive and negative media narrative indices. 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 interpretation index is associated with a 0.69% decline in quarterly fund flows into Chinese assets, while the estimated effect of the positive interpretation index is small and statistically insignificant. These results indicate that adverse media narratives play a dominant role in shaping cross-border institutional investment decisions.

Negative news apparently 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.

M.2 Higher Moment Effects of Media-Implied Risk.

Prediction 5 implies $\partial a/\partial \sigma^2_{i,t} < 0$, so heightened media-implied uncertainty reduces allocation independently of average sentiment. Hassan, Hollander, Van Lent and Tahoun (2019) also 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 domiciled in country c in month t (see Appendix G). 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 A9 column (8), after controlling for cross country macroeconomic and financial fundamentals, a one standard deviation increase in the risk per-

[38] 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.

ception 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.

Table A9: 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 (13) 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 M. 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.