When Words Move Money: Diplomatic Sentiment and International Capital Flows

Cunyi Yang[a], Shuchi Zhang[b], Ioannis Kyriakou[c,*], Nikos C. Papapostolou[c]

[a]*Faculty of Business and Economics, The University of Hong Kong*
[b]*School of Mathematics, Sun Yat-sen University*
[c]*Bayes Business School, City St George's, University of London*

*Corresponding author.

Email addresses: yangcy0126@gmail.com (Cunyi Yang), zhangshch27@mail2.sysu.edu.cn (Shuchi Zhang), ioannis.kyriakou@city.ac.uk (Ioannis Kyriakou), n.papapostolou-1@citystgeorges.ac.uk (Nikos C. Papapostolou)

When Words Move Money: Diplomatic Sentiment and International Capital Flows

Abstract

We construct a text-based measure of war-related diplomatic sentiment from 154,185 foreign-ministry communications across the 15 largest world economies. The daily index tracks military escalations and ceasefires, varies across countries, and predicts newspaper-based geopolitical risk more than the reverse. Adverse Chinese rhetoric foreshadows stronger southbound reallocation into Hong Kong equities and weaker Stock Connect flows; a one-unit decline shifts daily flows by $42.4 million towards outflows, operating through a relative-price channel widening the A-H premium rather than onshore declines. In monthly cross-country analyses, only the U.S. shows safe-haven behavior; adverse rhetoric raises Chinese and U.S. trading volume and U.S. volatility.
Keywords: diplomatic sentiment; capital flows; geopolitical risk

I. Introduction

Geopolitical tensions have become a persistent feature of the global economy. Great-power rivalry, recurrent regional conflicts, and the weaponisation of supply chains all force international investors to reassess geopolitical risk on a continuing basis (Caldara and Iacoviello, 2022). Financial markets respond to wars, sanctions, and smaller shifts in official language that change expectations about future conflict and cooperation (Pástor and Veronesi, 2012). Existing empirical measures, however, still treat international relations largely as discrete states – peace versus war or sanctions versus no sanctions – with limited ability to track continuously evolving diplomatic signals. We address this measurement gap by constructing a high-frequency index of war-related diplomatic sentiment from official foreign-policy communication and by studying how this measure relates to international capital flows.

Our point of departure is the large literature in international relations and political economy that seeks to measure the state of bilateral or multilateral relationships. Existing measures fall into three broad classes. One class uses discrete political events, such as formal alliances, war onsets, militarised interstate disputes, or sanctions episodes, often based on event datasets such as the Correlates of War project (Gowa, 1994; Hufbauer et al., 2007; Maoz, 2010). A second class uses revealed political behaviour, including UN voting alignment, diplomatic recognition, and the intensity of high-level visits and summits (Gartzke, 2006; Fuchs and Klann, 2013; Bailey et al., 2017). A third class, more common in macro-finance, uses news-based indices of geopolitical risk and policy uncertainty (Baker et al., 2016; Caldara and Iacoviello, 2022). Their common limitation is that they are typically low-frequency and highly discrete: they move mainly when conflict escalates into war, sanctions are imposed, or formal alignments change. They are useful for identifying large regime shifts but less well suited to measuring gradual changes in diplomatic tone or the more nuanced emotional content of official discourse.

Recent text-as-data work has begun to infer foreign-policy positions and diplomatic narratives from political speeches and official communications (Baturo et al., 2017; Mochtak and Turcsanyi, 2021). In finance, official policy documents have likewise been shown to contain economically meaningful sentiment: Correa et al. (2021) extract sentiment from central banks' financial stability reports and show that it helps explain the financial cycle and predict banking crises. Beyond academia, think tanks and private data providers have also begun to build NLP-based indicators of diplomatic engagement and conflict dynamics, such as HCSS's GINA Diplomatic and Permutable AI's conflict trackers (Kommandeur and van Dam, 2024; Permutable AI, 2024). Together, these efforts show that continuously updated geopolitical indicators are feasible. What remains limited, however, are high-frequency academic measures built directly from the day-to-day language of official diplomacy, transparently validated, and able to predict newspaper-based geopolitical risk before geopolitical conditions are fully absorbed by the news cycle. We address that gap by constructing daily country-level war-related diplomatic sentiment indices for the 15 largest world economies[1] from 154,185 official foreign-ministry and foreign-office texts using a multi-stage large language model (LLM) pipeline with built-in cross-validation and human audits, and by studying their implications for international equity flows.

We introduce the *War-related Diplomatic Sentiment Index* (WDSI), which measures the tone of a country's official statements on war, military tension, and security threats. Conceptually, diplomatic sentiment is a latent and continuous variable that summarises how friendly, neutral, or hostile a state sounds in its official communication at a given point in time. We infer this variation from the full text of foreign-policy communication at daily frequency. We treat war-related diplomatic sentiment as a structured text feature and extract it with modern language models. We then classify and score individual statements according to whether they convey concern, support, condemnation, or threats, and aggregate those scores into a daily index. The resulting WDSI provides a continuous high-frequency measure of official diplomatic tone that is well suited to temporal analysis and cross-country comparison.

We construct the WDSI across the 15 largest world economies in a unified official diplomatic-text database. The source material varies with institutional practice: the Chinese index is built from Ministry of Foreign Affairs press-conference transcripts; the U.S. index from Department of State press releases, briefings, remarks, readouts, announcements, and related official statements; and the indices for the remaining economies in the top-15 sample from comparable foreign-ministry releases, briefings, speeches, statements, and related official texts. The scoring architecture is common across countries, and the input unit follows the structure of each source.

[1] Countries with the largest nominal GDP worldwide in 2026 (in trillion U.S. dollars). Source: Statista, retrieved May 4, 2026, from https://www.statista.com/statistics/268173/countries-with-the-largest-gross-domestic-product-gdp/ (based on IMF World Economic Outlook Database).

Chinese MFA press conferences are split into question-answer exchanges; stand-alone releases, statements, remarks, readouts, and speeches enter as self-contained source records, with structured briefings split where the source text provides separable exchanges or sections. We then apply the same four steps: cleaning and boilerplate removal, war/security relevance screening, sentiment classification, and intensity scoring. The scored items are aggregated to the country-day level by publication date and then smoothed with a short moving average. The result is a continuous high-frequency index anchored in official diplomacy, separate from media coverage or realised-event coding, and comparable across countries within a unified measurement framework.

A natural concern is whether the index captures genuine variation in war-related diplomatic sentiment or mainly model artefacts. We therefore devote substantial attention to validation. First, we benchmark the WDSI against major geopolitical events, including military escalations and ceasefire agreements, and show that it moves in the expected direction and timing. Second, we compare it with the widely used geopolitical risk index. The WDSI exhibits only modest contemporaneous co-movement with GPR and retains substantial independent variation; in a monthly comparison with the China-specific GPR benchmark, months with especially adverse Chinese diplomatic sentiment show stronger and more persistent predictive power for future GPR than the reverse direction, consistent with policy-signal information and limited duplication. Third, for the Chinese MFA corpus that anchors the main high-frequency analysis, we conduct manual audits with coders trained in international relations to assess whether the automated labels align with independent qualitative readings of the same Q&A units. Finally, we show that the WDSI predicts outcomes for which war-related sentiment should matter, including cross-border capital flows, stock market returns, and market activity, even after controlling for standard macro-financial variables.

Our main empirical tests focus on international capital flows, with China as the primary high-frequency setting because Stock Connect reports daily, directionally decomposed cross-border equity flows. China's capital-account management routes these transactions through clearly defined programs, creating a rare setting in which daily international flows are both observable and decomposable by direction. We can therefore distinguish offshore purchases of A-shares (northbound) from mainland purchases of Hong Kong equities (southbound) and map changes in diplomatic tone to high-frequency reallocations in a tightly identified trading channel. For most other large economies, public international-flow measures are monthly or lower-frequency, more aggregated, and less closely tied to bilateral trading interfaces, which weakens comparability and attenuates short-horizon inference. More broadly, our focus on cross-border capital reallocation is consistent with evidence that political uncertainty reshapes other forms of international investment, including cross-border acquisitions (Cao et al., 2019).

Accordingly, the empirical analysis has three layers, each dictated by the frequency and informativeness of the available flow data. The first layer uses Chinese Stock Connect as the daily flow laboratory and relates both Chinese and foreign WDSIs to inflow, outflow, and net flow. The second layer assembles monthly own-country international-equity measures wherever official data permit, yielding a core monthly sample for the United States, Japan, South Korea, Brazil, and Canada, together with monthly proxy evidence for France, Italy, and Russia. The third layer summarises these monthly cross-country linkages in pairwise sentiment-flow networks that visualise directional spillovers in a six-country core system. We also examine stock returns, trading volume, return volatility, and the A-H premium[2] to assess whether war-related

[2]The A-H premium is defined as the percentage price difference between mainland A-shares and their corre-

sentiment affects portfolio reallocation, trading intensity, relative pricing, and risk absorption. Throughout, our empirical framework relies on time-series regressions with a broad control set for global and domestic financial conditions, including equity returns, geopolitical risk, economic policy uncertainty, exchange rates, and safe-haven demand.

A central design choice is to treat cross-border capital flows as the primary dependent variable and stock returns as a complementary outcome. In semi-strong efficient markets, publicly available text signals can be incorporated into prices quickly, so close-to-close return predictability is often weak even when the underlying information is economically important. Text-based signals are frequently more informative for quantity and risk outcomes, such as trading activity, volatility, and portfolio reallocation, than for aggregate daily returns. This emphasis on quantity-based outcomes is consistent with recent evidence from China showing that foreign order flow predicts future returns and news at the firm level (Lundblad et al., 2026). It is also consistent with international-finance evidence that treats equity flows as first-order objects with their own persistence and adjustment dynamics (Froot et al., 2001; Brennan and Cao, 1997; Portes and Rey, 2005; Griffin et al., 2004; Hau and Rey, 2006). To complement the flow-based baseline, Section IV presents additional price-side evidence from Chinese equity-index returns under matched control specifications.

The main findings fall into two broad categories: measurement and economic impact. On the measurement side, the 15-economy WDSI database provides a fine-grained representation of war-related diplomacy. It reacts promptly to major security events, distinguishes between escalation and de-escalation, and captures gradual shifts in diplomatic tone that are largely invisible

sponding H-shares listed in Hong Kong (with A-share prices converted into Hong Kong dollars). A higher value indicates that A-shares trade at a larger premium over H-shares.

in standard discrete indicators. Relative to war dummies, sanctions episodes, and newspaper-based geopolitical risk measures, it offers both higher temporal resolution and stronger predictive information about subsequent realisations of such risk. In the China benchmark, monthly low-point WDSI predicts future China-specific GPR over a wider lag range than the reverse direction. On the economic side, more negative Chinese war-related diplomatic sentiment is associated with stronger southbound reallocation and weaker net Stock Connect flows: in the baseline daily specification, a one-unit decline in $WDSI^{CN}$ lowers net Stock Connect flows by \$42.4 million. The mechanism works through relative pricing: adverse rhetoric widens the A-H premium, making Hong Kong-listed shares more attractive relative to mainland-listed shares. In the daily foreign-country spillover regressions, many foreign WDSIs load most strongly on the Chinese outflow equation, indicating that outward reallocation through the southbound channel is the dominant cross-country margin. Among countries with sufficiently informative monthly own-country capital-flow data, the United States is the only core market that displays a clear safe-haven pattern. South Korea moves in the opposite direction, with more adverse sentiment associated with weaker inflows or net positions, while estimates for Canada, Japan, and Brazil are statistically weak. The six-country pairwise network shows concentrated directional transmission, with the United States and South Korea as the main receiving nodes and Brazil and Canada contributing economically meaningful bilateral links. Taken together, the results suggest that only the strongest safe-haven markets can reliably convert adverse domestic war-related rhetoric into capital inflows; other markets either lose capital or receive less inflow, while the remaining estimates are statistically weak. This asymmetric redistribution of portfolio demand provides a capital-flow perspective on geopolitical risk: adverse diplomatic rhetoric serves as a risk signal for prices, reallocates international capital, and reveals the economic stakes that may shape geopolitical-risk episodes. Additional price-side evidence argues against a simple absolute-price-decline channel. Daily Chinese return responses are weak, while the A-H premium, trading activity, and medium-horizon repricing remain informative.

This paper contributes to three strands of literature. First, it contributes to international relations and international political economy by moving beyond discrete conflict and alliance indicators to a continuous, text-based measure of diplomatic sentiment. By translating official diplomatic language into a time-varying index, we provide a tool for studying the dynamics of interstate relations at much higher frequency and granularity than traditional event data; see Baum and Zhukov (2019) and Gentzkow et al. (2019) for related efforts to quantify foreign-policy stances from text. Second, building on the literature on political risk and international investment (Papaioannou, 2009; Julio and Yook, 2012; Kelly et al., 2016; Cao et al., 2019), we contribute to international macro-finance by introducing a country-level political-sentiment measure that can be linked to trading, investment, and portfolio allocation. Our results show that war-related diplomatic sentiment helps explain cross-border capital flows and market outcomes (Benhima and Cordonier, 2022; Feng et al., 2023; Choi and Havel, 2025) even after controlling for conventional measures of geopolitical risk and macroeconomic conditions. By showing that adverse diplomatic rhetoric reallocates portfolio demand across markets, the paper also offers a capital-flow perspective on the economic interests and potential incentives surrounding geopolitical risk. Third, we contribute to the growing literature on text-based measurement in economics and finance by showing how large language models can be used to construct structured indices of "soft" variables such as sentiment, hostility, and trust from official documents (Buehlmaier and Whited, 2018; Gentzkow et al., 2019; Hassan et al., 2019; Correa et al., 2021). More broadly, the paper provides a general framework for converting qualitative diplomatic communication into quantitative data that can be used in empirical work.

The rest of the paper proceeds as follows. Section II introduces the concept of war-related diplomatic sentiment, describes the 15-economy official text database, and details the common construction of the WDSI. Section III presents the main empirical design and the baseline results for Chinese cross-border equity flows. Section IV reports 14 foreign-country spillovers into Chinese flows, monthly own-country international-equity evidence, pairwise network spillovers, and additional price-side and market-response evidence, including the A-H share premium, complementary return evidence, and market activity. Section V concludes.

II. Construction of the War-Related Diplomatic Sentiment Index

The construction of the WDSI proceeds in four steps: definition, measurement, audit, and presentation. The measurement sample is a unified official diplomatic-text database for the 15 largest world economies. We begin by defining war-related diplomatic sentiment, then describe the common measurement architecture applied to all country corpora, use selected country-level and cross-country displays to make the index transparent, assess reliability through structured validation exercises, and finally evaluate whether the index moves in ways that are consistent with observable security events and external risk benchmarks.

A. Diplomatic Sentiment and War-Related Diplomatic Sentiment

Diplomatic sentiment, as used in this paper, captures the emotional tone and attitudinal orientation conveyed in a state's official statements about international events, the conduct of other nations, and the state of bilateral or multilateral relations. Unlike personal emotion, it is expressed through formal and strategic language in channels such as foreign-ministry briefings and official remarks. We therefore treat it as a policy-relevant soft signal that reveals national preferences, risk tolerance, and willingness to escalate or de-escalate a situation.

On this basis, we define war-related diplomatic sentiment as the tone and evaluative stance embedded in a state's official foreign-policy discourse when it addresses war, armed conflict, military confrontation, security crises, or the possible onset, escalation, or easing of such situations. Unlike broader notions of geopolitical risk or general security tension, this concept captures both the occurrence of conflict-related events and the way a state publicly characterises the use of force and its anticipated consequences. Our focus therefore combines the presence of war or conflict with the tone through which the foreign ministry signals its position.

In terms of scope, war-related diplomatic sentiment covers at least three classes of statements. The first concerns active or highly salient armed conflicts, such as regional wars, invasions, and large-scale military operations. In this setting, expressions of concern, sympathy, solidarity, or support for ceasefire and negotiation all fall within scope. The second concerns latent military frictions and security risks, including joint exercises, arms sales, alliances, border or maritime disputes, nuclear and missile programs, and changes in defence treaties or base deployments. Here, official warnings, objections, firm positions, and conciliatory signals likewise convey assessments of war risk. The third concerns political events that may trigger military escalation, such as coups, violent clashes, terrorist attacks, or the breakdown of major security treaties. Statements framed in terms of international security or armed confrontation are therefore included even when large-scale fighting has not yet begun. Issues confined to economic, cultural, or technological affairs fall outside the definition unless they contain an explicit security or military dimension.

B. Measurement: Data, Methodology, and Index Construction

B.1. Data

The WDSI database covers the 15 largest world economies and is built from official diplomatic texts collected from foreign-ministry or foreign-office sources. Online Appendix II lists the country corpora, source institutions, coverage windows, and document counts. For the Chinese corpus, $WDSI^{CN}$ is constructed from transcripts of regular press conferences of the Ministry of Foreign Affairs from 5 January 2010 to 17 April 2026; all transcripts released on the same day are merged into a single daily observation. The same scoring scale, daily alignment principle, and smoothing rule are applied to all country corpora, subject only to source-specific document formats and coverage windows. The processed country-level WDSI series used in the empirical analysis are publicly available on a companion website (`https://yangcy0126.github.io/dsi-index-site/`). The WDSI dataset is maintained as a live research database, with periodic updates to coverage windows and documented country corpora.

For the Chinese corpus, regular MFA press conferences are the relevant official source for both institutional and econometric reasons. Institutionally, these briefings represent the central foreign-policy authority and are immediately amplified through official and mainstream media, giving them both formal status and broad visibility. Substantively, the limited number of questions in each briefing concentrates attention on the most salient war, conflict, and security developments of the day, so the resulting discourse is closely tied to contemporaneous shocks and directly addresses the use of force and its risks. Econometrically, questions are largely selected by journalists, whereas the spokesperson's responses are typically pre-cleared within the bureaucracy. Because briefings usually occur before the domestic market closes, short-run financial conditions are less likely to feed back into the tone of the statements. Relative to indices built from ex post news coverage, this source structure more closely approximates a policy signal.

B.2. Methodology: LLM-Based Multi-Stage Sentiment Extraction

Given our measurement objective and the structure of the data, we use an evaluation pipeline that combines large language models with task-specific procedures. The pipeline decomposes sentiment assignment into a sequence of simpler inference tasks. This differs from the horizontal decomposition common in existing work, where several related sub-indices are later aggregated into a composite measure. Our approach is vertical: a single sentiment label is produced through successive stages, each handled by an AI agent designed for a specific subtask. Relative to one-shot scoring, this vertically decomposed multi-agent design improves reliability and interpretability by simplifying each inference step and adding a consensus check, consistent with the broader text-as-data measurement literature (Buehlmaier and Whited, 2018; Gentzkow et al., 2019; Hassan et al., 2019).[3]

At each stage, we also impose cross-validation. Multiple independent judgements are generated, and the pipeline advances only when they agree. This rule makes the classification procedure more transparent and reduces the influence of idiosyncratic model errors. Relative to existing LLM-based sentiment applications, the resulting framework offers greater interpretability and robustness while remaining flexible enough to accommodate heterogeneous linguistic styles and multiple layers of sentiment.

Detailed procedures and implementation results are reported in Online Appendix III, while Figure 1 provides a simplified schematic. The pipeline begins by screening source-specific text

[3]Consistent with this rationale, Online Appendix IV.C reports that the multi-step pipeline outperforms single-step direct scoring by about 5-10 percentage points in overall accuracy and reduces occasional reasoning errors.

passages for war/security relevance, then classifies the sentiment direction of relevant passages, and finally applies custom-designed AI agents to quantify sentiment intensity. These steps together produce the final index.

[Figure 1 about here.]

B.3. Index Construction

Building on the procedure above, each source-specific text unit receives a discrete sentiment score on the [-3, 3] scale, with lower values indicating more negative sentiment. Country-day observations are formed by publication date: scored items released on the same day are aggregated within country-day, non-publication days inherit the most recent observed official stance, and the resulting daily series is smoothed with the same moving-average rule. This produces a comparable daily panel for all 15 economies while preserving the source-specific text structure used at the scoring stage. Table 1 reports the distribution of scores and several illustrative examples for the Chinese Q&A implementation, which provides a transparent view of how the common scoring rule operates at the text-unit level.

[Table 1 about here.]

To reduce the noise in single-day readings, we smooth each initial daily series with a seven-day rolling average. Figure 2 reports the Chinese country-level implementation and also shows its annual rolling mean and volatility. Over time, volatility declines while the mean drifts downwards, consistent with increasingly negative Chinese diplomatic rhetoric in the military domain.

[Figure 2 about here.]

C. Audit: Validity and Reliability Verification

Text-based indicators naturally raise two concerns. The first is whether the model correctly identifies war- or military-related content and assigns its sentiment accurately. The second is whether the resulting index is overly sensitive to technical choices and therefore contaminated by systematic measurement error. We address these concerns through three sets of validation exercises: manual auditing, model comparison and selection, and robustness checks.

First, we conduct a manual audit of the Chinese MFA press-conference corpus, which is the source of the main daily WDSI used in the high-frequency flow analysis. We draw a stratified random subset of about 1,200 Chinese MFA question-answer units at evenly spaced monthly intervals. Two research assistants trained in international relations independently code these materials without access to the model output. They record (i) whether the Q&A unit is war- or military-related, (ii) the sentiment category, and (iii) sentiment intensity. Inter-annotator agreement is about 88 percent for the war- or military-related indicator, 80 percent for sentiment category, and 83 percent for sentiment intensity. This audit validates the core Chinese Q&A setting; cross-country comparability is assessed through the common extraction pipeline and the event, GPR, and empirical validation exercises below.

Second, we benchmark large language models against more traditional methods. Before formally constructing the WDSI, we evaluate three method families on the manually labelled sample described above: (i) hand-crafted dictionaries and keyword counts, (ii) shallow supervised models built on word embeddings, and (iii) several large language models that can process mixed Chinese and English input. The Qwen-3 model used in this paper outperforms the alternatives at each stage of the pipeline.

Third, we compare a single-step direct-scoring strategy with the multi-step decomposition used in our workflow, in which multiple agents must agree before the process moves forward. For both human coders and AI agents, the latter approach materially reduces error rates by lowering the cognitive burden at each individual step.

Detailed procedures and summary statistics for these validation exercises are reported in Online Appendix IV.

D. Understanding the War-Related Diplomatic Sentiment Index

D.1. Event Consistency and Cross-Country Patterns

We next examine whether the WDSI behaves like an interpretable measure of official war-related policy tone. The goal of this section is descriptive validation: the index should move with recognisable security episodes, display country-specific variation, avoid collapsing into a mechanically common global factor, and retain information that is distinct from newspaper-based geopolitical-risk measures. Figures 2 and 3 present two event-rich country-level panels from the same 15-economy WDSI database. The U.S. textual corpus is drawn from Department of State press releases, briefings, remarks, readouts, announcements, and related official statements from 1 January 2010 to 18 April 2026; a detailed description of the data is provided in Online Appendix II. Figure 3 reports the U.S. series, $WDSI^{US}$.

[Figure 3 about here.]

Figure 4 presents the full 15-economy WDSI database on a common scale. The panels show substantial cross-country heterogeneity in both the level and persistence of war-related diplomatic rhetoric. Some countries, such as China, the United States, the United Kingdom, Russia, and South Korea, display sustained negative phases, whereas others move closer to a neutral baseline for long stretches and exhibit sharper episode-specific swings. This heterogeneity is useful for the empirical analysis below because it rules out a single global news factor mechanically shared across countries. The index provides country-specific policy-signal variation that can be linked to cross-border portfolio flows and market responses.

[Figure 4 about here.]

Pairwise co-movement provides the same message. If the WDSI merely captured a common global news factor, major countries' series would move closely together. The data show only weak overall co-movement between $WDSI^{CN}$ and $WDSI^{US}$, with a correlation coefficient of 0.09. This low correlation is consistent with the absence of sustained direct military confrontation between China and the United States and with the fact that the two series load on different theatres. $WDSI^{CN}$ is more closely tied to rhetoric on East Asia and China's immediate periphery, especially references to Japan, the Republic of Korea, India, and cross-Strait issues, whereas $WDSI^{US}$ is more responsive to statements on Middle Eastern conflicts and Russia. The correlation rises to 0.18 in 2016, during the South China Sea stand-off, but falls to -0.17 after Donald Trump returned to office on 20 January 2025. This reversal is consistent with a period of intensified Sino-American rivalry, a more restrained U.S. military posture, and more assertive Chinese official rhetoric. More broadly, the cross-country panels show that the index responds to each government's own security environment and official policy stance without imposing a uniform global sentiment path.

D.2. Comparison with Geopolitical Risk

The closest and most widely used benchmark in the literature is the Geopolitical Risk Index (GPR) of Caldara and Iacoviello (2022). Although both the GPR and the WDSI focus on war, conflict, and security, they capture distinct layers of geopolitical information.

The GPR is essentially a newspaper-based event-risk measure. It tracks reports of objective developments such as war threats, terrorist attacks, and military confrontations, and quantifies the frequency and intensity of geopolitical events together with the uncertainty they generate. The WDSI is a sentiment-based measure extracted from official diplomatic communication. It captures the tone and policy stance embedded in government statements regarding war and the use of force. Put differently, the GPR is closer to the realised salience of geopolitical events in the news, whereas the WDSI is closer to the evolving policy signal conveyed through official rhetoric. The two indices are related and contain distinct information.

This distinction also follows from how the two measures are constructed. The GPR is built from newspaper coverage and therefore reflects media attention and the news cycle. The WDSI is derived from routine press conferences and official statements issued by central foreign-policy authorities. As a result, the WDSI can move when official rhetoric hardens or softens even when geopolitical developments have limited visibility in media-based risk measures. The GPR tends to react more strongly once conflicts become salient as realised events. This timing difference suggests that the WDSI should be especially informative for predicting subsequent GPR movements when official rhetoric reaches unusually adverse levels.

In the data, the two indices exhibit only modest contemporaneous co-movement. The correlation between China's baseline WDSI and the global GPR index is about -0.16. The corresponding correlations with the China- and U.S.-specific GPR indices are about -0.17 and -0.12, respectively. These relatively low contemporaneous correlations indicate that the WDSI contains information beyond existing GPR measures. At the same time, weak contemporaneous co-movement remains compatible with a lead-lag structure in which diplomatic sentiment predicts later news-based risk salience more strongly than it is predicted by current GPR. In particular, the GPR tends to spike more sharply around realised conflict events, whereas the WDSI moves as a measure of official policy tone that can turn before the risk is fully reflected in the news.

To examine this timing channel directly, we estimate monthly bivariate VARs between a low-point measure of Chinese diplomatic sentiment and the China-specific geopolitical risk index GPR_m^{CN}. The sentiment measure, denoted $WDSI_m^{CN,\min}$, is the minimum raw daily Chinese WDSI value within month m. This specification asks whether months containing especially adverse official war-related rhetoric predict subsequent newspaper-based geopolitical risk and avoids mechanically repeating a monthly GPR value across daily observations. We estimate VAR(p) systems over $p = 1, \ldots, 12$ months and conduct Granger-causality tests in both directions at each horizon. Table 2 reports the full profile of the resulting p-values. The evidence shows stronger predictive content from adverse diplomatic sentiment to subsequent geopolitical-risk salience than in the reverse direction. The null that WDSI does not predict GPR is rejected at the 5% level for lags 1–6 and remains marginal at lags 7–9. The reverse-direction null is rejected only at lags 2–3. The BIC selects a two-month lag, where both directions are statistically significant, indicating short-run feedback; however, the WDSI$\rightarrow$GPR pattern is stronger and more persistent across the lag grid. The sum of the lagged WDSI coefficients in the GPR equation is negative throughout, so lower diplomatic sentiment predicts higher subsequent GPR, as expected.

[Table 2 about here.]

This monthly lead-lag structure matters for both measurement and empirical use. Relative to the monthly GPR benchmark, the WDSI low-point measure provides stronger forward-looking policy-signal variation about subsequent geopolitical-risk salience than the reverse GPR-to-WDSI direction. For investors and other forward-looking agents, this predictive signal means that especially adverse official rhetoric can help identify subsequent geopolitical deterioration and the repricing and portfolio adjustment that may follow. The WDSI adds value by predicting future news-based risk salience with official policy-signal information that contemporaneous media-based risk measures do not fully incorporate.

III. War-related Diplomatic Sentiment and Capital Flows

A. Baseline Results: The Case of China

We now turn to the first empirical layer: China's high-frequency flow setting. We examine how war-related diplomatic sentiment affects cross-border equity flows through the Stock Connect programs, which record daily trading in Chinese equities between mainland Chinese and international investors through both the northbound and southbound channels[4].

Our baseline specification relates daily capital flows to war-related diplomatic sentiment while controlling for domestic and global financial conditions. Following He et al. (2023), we estimate for the Chinese market

$$Flow_{k,t}^{CN} = \alpha_k^{CN} + \beta_k^{CN} WDSI_t^{CN} + \boldsymbol{\gamma}_k^{CN\prime}\mathbf{R}_t + \boldsymbol{\delta}_k^{CN\prime}\mathbf{Z}_t + \mu_{m(t)} + \varepsilon_{k,t}^{CN}\,, \tag{1}$$

where t indexes trading days and $k \in \{\text{Inflow}, \text{Outflow}, \text{Net flow}\}$. The dependent variable $Flow_{k,t}^{CN}$ is constructed from Stock Connect trading. Inflow is net northbound trading volume

[4]The Stock Connect schemes link the Shanghai and Shenzhen exchanges with the Hong Kong Stock Exchange through two distinct channels. Northbound trading allows overseas and Hong Kong-based investors to trade eligible A-shares listed on the mainland exchanges, whereas southbound trading allows mainland Chinese investors to trade eligible Hong Kong-listed equities. Accordingly, our daily flow measures capture the joint behaviour of foreign investors (primarily in the northbound segment) and domestic investors reallocating wealth into offshore Chinese listings (through the southbound segment).

and captures net purchases of mainland A-shares through the northbound channel. Outflow is net southbound trading volume and captures net purchases of Hong Kong-listed equities through the southbound channel. Net flow combines northbound and southbound trades into an aggregate measure of net cross-border capital movement associated with Stock Connect activity. All three series are winsorised at the 5th and 95th percentiles in levels (He et al., 2023).

The key explanatory variable, $WDSI_t^{CN}$, is the seven-day moving average of the Chinese war-related diplomatic sentiment index constructed from Ministry of Foreign Affairs press-conference transcripts using a large language model, as described in Section II. Higher values correspond to more positive war-related diplomatic sentiment. Accordingly, a deterioration in diplomatic tone is captured as a decline in $WDSI_t^{CN}$.

The control vector $\mathbf{R}_t$ includes three equity return series: the United States (CRSP Ret), developed markets excluding the United States (Developed ex US Ret), and China (A-share Ret).[5] These controls follow the international finance literature showing that cross-border equity flows co-move strongly with both domestic and global market performance (e.g., Froot et al., 2001; Griffin et al., 2004; Portes and Rey, 2005; He et al., 2023). The control vector $\mathbf{Z}_t$ includes the Chinese and U.S. geopolitical risk indices[6] (Caldara and Iacoviello, 2022), the Chinese and U.S. economic policy uncertainty indices[7] (Baker et al., 2016), gold returns (defined

[5]The U.S. and developed-market equity returns are obtained from the data library of Kenneth R. French: `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french`. The Chinese equity market return is from the CSMAR database. The developed-market return corresponds to a broad developed-market portfolio that excludes the United States, so as to avoid mechanically double counting U.S. market movements when both R^{US} and the developed-market return are included as controls. All three market-return series are winsorised at the 5th and 95th percentiles to mitigate the influence of extreme market conditions (e.g., the global outbreak of COVID-19 in March 2020) (He et al., 2023).

[6]Data source: `https://www.matteoiacoviello.com/gpr.htm`.

[7]Data source: `https://www.policyuncertainty.com/index.html`.

as log first differences of gold prices), which capture short-run adjustments in safe-haven asset dynamics (Baur and McDermott, 2010), and the bilateral RMB-USD exchange rate, which is tightly linked to international portfolio rebalancing (Caporale et al., 2015; Gelman et al., 2015). Monthly GPR and EPU series are mapped to daily frequency by assigning the monthly value to each trading day within the month. All control variables are winsorised at the 5th and 95th percentiles. Month-of-year fixed effects $\mu_{m(t)}$, where $m(t) \in \{1, \ldots, 12\}$ denotes the calendar month of the year, absorb recurring within-year seasonality. They therefore leave month-specific shocks in a given year available for identification, and the monthly GPR and EPU controls remain identified from their variation across months over time. The error term $\varepsilon_{k,t}^{CN}$ captures idiosyncratic shocks. Online Appendix Table I.1 provides compact definitions of the variables used across the empirical analysis. We estimate Equation (1) separately for the three flow measures using ordinary least squares with heteroskedasticity-robust standard errors. Table 3 reports descriptive statistics for all variables used in the baseline regressions.

[Table 3 about here.]

Table 4 reports the baseline results for daily Stock Connect flows. Columns (1)–(3) correspond to net northbound inflows, net southbound flows, and their combined net flow. The coefficient on Chinese war-related diplomatic sentiment, $WDSI_t^{CN}$, is 0.009 and statistically insignificant in the inflow regression, indicating little response in net northbound purchases of A-shares by overseas and Hong Kong-based investors. The outflow regression shows a different pattern: the coefficient is -0.299 and highly significant, indicating that a deterioration in China's war-related diplomatic tone is associated with larger net southbound purchases of Hong Kong-listed equities by mainland investors and, therefore, larger capital outflows from the mainland through Stock Connect.

[Table 4 about here.]

The net-flow regression in column (3) yields a positive and statistically significant coefficient of 0.424 on $WDSI_t^{CN}$. Because net flow is defined as net northbound purchases minus net southbound purchases and the inflow response in column (1) is economically small and statistically insignificant, the net-flow effect is driven almost entirely by the southbound margin. The coefficient in the net-flow equation should nevertheless be read directly from column (3), since the three equations use slightly different nonmissing samples (2,205, 2,224, and 2,190 observations, respectively), and each flow measure is winsorised separately after construction. More positive war-related sentiment is therefore associated with smaller net capital outflows from the mainland, whereas a deterioration in diplomatic tone shifts Stock Connect activity towards net outflow. Economically, a one-unit decline in the seven-day moving average of $WDSI_t^{CN}$ is associated with a decline of approximately \$42.4 million in daily net Stock Connect flows.

Consistent with a large international finance literature showing that cross-border equity flows are strongly procyclical and closely linked to local equity returns through return-chasing and portfolio-rebalancing motives (e.g., Bohn and Tesar, 1996; Brennan and Cao, 1997; Froot et al., 2001; Griffin et al., 2004), the Chinese market return enters positively and very precisely in all three flow regressions. Higher A-share returns are associated with stronger northbound inflows and stronger southbound trading, indicating procyclical position adjustments on both sides of Stock Connect. More broadly, this pattern is consistent with evidence that international equity transactions are shaped by expected returns, information frictions, and other cross-border impediments (Portes and Rey, 2005).

The coefficients on the U.S. and developed-market return controls also indicate systematic rebalancing across major markets. Because net flow is defined as northbound inflow minus southbound outflow, any regressor that increases southbound outflows lowers net flows unless it is offset by a sufficiently strong inflow response. In line with this accounting, higher U.S. market returns are associated with larger southbound outflows and, together with a modest decline in northbound inflows, with lower net flows. Higher returns in the rest of developed markets (excluding the United States) are associated with stronger northbound inflows and weaker southbound outflows, implying higher net flows.

The RMB-USD exchange rate also enters with large and statistically significant coefficients: RMB depreciation is associated with weaker northbound inflows and stronger southbound outflows, consistent with the role of exchange-rate movements and currency risk in cross-border portfolio adjustment (Hau and Rey, 2006; Camanho et al., 2022). Geopolitical risk exhibits a more differentiated pattern. Higher Chinese geopolitical risk is associated with stronger northbound inflows, weaker southbound outflows, and higher net flows, whereas higher U.S. geopolitical risk is mainly linked to weaker northbound inflows and lower net flows. Policy uncertainty and gold returns are included to proxy for broader risk conditions and short-run adjustments in safe-haven demand (Baker et al., 2016; Baur and McDermott, 2010). In the baseline specification with month-of-year fixed effects, the China EPU index and gold returns are not precisely estimated, whereas the U.S. EPU index remains statistically significant.

B. Robustness Checks

We conduct two robustness checks to verify that the baseline results for Chinese Stock Connect flows are robust to temporal aggregation and dynamic feedback. First, we vary the smoothing window used to aggregate the Chinese sentiment series and show that the baseline results remain stable (Section III.B.1). Second, we use a residualised VAR framework to assess the dynamic consistency of the sentiment-flow relation (Section III.B.2). Supplementary inference checks based on HAC and moving-block bootstrap procedures are reported in Online Appendix IV.E. We address non-war sentiment placebos and other identification concerns in Section III.C.

B.1. Smoothing-Window Robustness

Our baseline Chinese sentiment index, $WDSI_t^{CN}$, uses a seven-day moving average of daily war-related diplomatic sentiment. To assess whether the results depend on this smoothing choice, we re-estimate Equation (1) using a three-day average, $WDSI_t^{CN,3}$, and a thirty-day average, $WDSI_t^{CN,30}$.

Table 5 reports the results. Columns (1)–(3) replace the baseline seven-day average with the three-day measure, and columns (4)–(6) use the thirty-day measure. The qualitative pattern is unchanged in both cases. With $WDSI_t^{CN,3}$, the coefficient remains insignificant in the inflow regression, negative and highly significant in the outflow regression, and positive and significant in the net-flow regression. With $WDSI_t^{CN,30}$, the same sign pattern obtains, but the magnitudes are larger: the outflow and net-flow coefficients rise to -0.669 and 0.977, both significant at the 1% level, while the inflow coefficient remains insignificant. Overall, the three-day specification yields slightly smaller estimates than the seven-day benchmark, whereas the thirty-day specification produces materially larger effects, consistent with more persistent deteriorations in war-related diplomatic sentiment being associated with larger cross-border outflows from China.

[Table 5 about here.]

These estimates reinforce the baseline conclusion that Chinese war-related diplomatic sentiment predicts cross-border equity flows and that this relation is robust to reasonable changes in temporal aggregation. Whether sentiment is measured over a short three-day window or a more persistent thirty-day window, more negative war-related rhetoric continues to predict stronger southbound outflows and weaker net capital inflows.

B.2. VAR-Based Dynamic Consistency Check

To complement the OLS evidence, we estimate residualised bivariate VAR models to assess whether the sentiment-flow relation survives dynamic feedback. Following Equation (1), we first partial out the full control block – China, U.S., and developed-market returns; China and U.S. GPR and EPU; gold; FX; and month-of-year fixed effects – from both sentiment and each flow channel, and then estimate VARs on the residual series.

To keep the setup concise and directly comparable, we report four residualised VAR(1) specifications. Panels A and B use daily data with $WDSI_t^{CN}$ and $WDSI_t^{CN,30}$, respectively. Panels C and D provide weekly counterparts constructed from Monday-anchored weeks, with flow variables summed within week and sentiment and controls averaged. Table 6 reports results for all three flow channels – inflow, outflow, and net flow – together with the associated Granger-causality tests.

The channel pattern is consistent across specifications. In the daily VARs, sentiment significantly predicts outflow under both sentiment definitions and predicts net flow under the thirty-day index. The weekly VARs deliver the same directional pattern. Under weekly $WDSI_t^{CN,30}$, the sentiment coefficient is significant for outflow (-0.079, p-value = 0.040) and net flow (0.105, p-value = 0.011), while inflow remains insignificant; under weekly $WDSI_t^{CN}$, outflow remains significant (-0.091, p-value = 0.016) and net flow is marginally significant (0.080, p-value = 0.052). Figure 5 reports impulse responses with 95% confidence bands for all four specifications and confirms the same dynamic ordering across flow channels.

[Table 6 about here.]

[Figure 5 about here.]

C. Addressing Endogeneity Concerns

Our baseline specification in Equation (1) treats war-related diplomatic sentiment as exogenous to cross-border equity flows. In practice, one measurement concern and three endogeneity concerns remain. The measurement concern is that generic diplomatic tone may be the source of the estimates. The endogeneity concerns are that Stock Connect flows may feed back into the tone of official diplomacy, both rhetoric and flows may respond to common slow-moving forces, and the sentiment measure may still contain residual noise correlated with omitted determinants of capital flows.

To address these concerns, we implement four complementary checks. First, we run non-war sentiment placebos to test whether the Stock Connect results reflect generic diplomatic tone or war-related rhetoric. Second, we use future sentiment and detrended sentiment specifications to assess reverse causality and low-frequency co-movement. Third, we estimate dynamic specifications with lagged flows and controls to absorb serial correlation and slow-moving co-movement. Fourth, we use an instrumental-variables design that exploits cross-spokesperson variation in average diplomatic tone as a source of exogenous shifts in the sentiment index.

C.1. Non-War Sentiment Placebos

A natural concern is that the baseline results may be capturing a generic response to negative diplomatic language. We therefore re-estimate Equation (1) using alternative sentiment measures extracted from the same MFA press-conference corpus in place of $WDSI_t^{CN}$. Specifically, we construct (i) an economic-related sentiment index, $EDSI_t^{CN}$, capturing rhetoric related to economic and trade disputes, and (ii) an index of other diplomatic sentiment, $ODSI_t^{CN}$, capturing non-military and non-economic issues such as human rights, climate, public health, consular affairs, and cultural or educational exchanges. If the baseline relation simply reflects generic negative official language, these non-war sentiment indices should generate similar flow responses. We also consider a truncated version of the war-related index that isolates negative episodes only. Specifically, we set $WDSI_t^{CN,-}$ equal to $WDSI_t^{CN}$ when the daily index is negative and to zero otherwise, and then re-estimate Equation (1) using $WDSI_t^{CN,-}$ as the main regressor.

Table 7 reports the nine regressions implied by this placebo exercise. Columns (1)–(3) use the economic-conflict sentiment index, columns (4)–(6) use the other diplomatic sentiment index, and columns (7)–(9) use the negative-part war index.

[Table 7 about here.]

For the economic-conflict sentiment in columns (1)–(3), the coefficients are statistically insignificant across all three flow specifications. The estimates therefore provide no evidence of a systematic relation between economic-conflict rhetoric and Stock Connect flows.

A similar pattern obtains for the other diplomatic sentiment index in columns (4)–(6). Although the inflow and net-flow coefficients remain negative, none of the three estimates is statistically distinguishable from zero. Thus, once the sentiment measure moves away from war-related rhetoric, the baseline flow pattern no longer survives.

Columns (7)–(9), which use the negative-part war index $WDSI_t^{CN,-}$, reproduce the qualitative conclusion of the baseline results. The coefficient on $WDSI_t^{CN,-}$ remains statistically insignificant in the inflow specification and is negative (positive) and statistically significant in the outflow (net-flow) specification. If anything, the net-flow coefficient is somewhat stronger than in the baseline specification. Because $WDSI_t^{CN,-}$ varies only on days with negative war-related sentiment and becomes more negative as rhetoric turns more hawkish, the positive net-flow coefficient implies that adverse war-related rhetoric continues to predict lower net capital inflows, or equivalently larger net outflows, even when positive war-related statements are excluded altogether.

Overall, the placebo exercise in Table 7 supports the interpretation that Chinese Stock Connect flows respond specifically to war-related diplomatic sentiment, with little evidence of a comparable response to generic negative rhetoric or other types of diplomatic language.

C.2. Reverse-Causality and Low-Frequency Trend Checks

We next test for reverse causality by asking whether Stock Connect flows predict future war-related diplomatic sentiment. Specifically, we estimate reverse-causality regressions in which the dependent variable is the one-day-ahead sentiment index, $WDSI_{t+1}^{CN}$, and the key regressors are contemporaneous and lagged flows:

$$WDSI_{t+1}^{CN} = \alpha + \rho_0 Flow_{k,t}^{CN} + \rho_1 Flow_{k,t-1}^{CN} + \boldsymbol{\gamma}'\mathbf{R}_t + \boldsymbol{\delta}'\mathbf{Z}_t + \mu_{m(t)} + u_{k,t}, \tag{2}$$

for $k \in \{\text{Inflow}, \text{Outflow}, \text{Net flow}\}$, where $\mathbf{R}_t$ and $\mathbf{Z}_t$ denote the same control sets as in the baseline regressions, and $u_{k,t}$ is an error term. Because Stock Connect flows are observed only on trading days whereas the WDSI is a calendar-day series, $Flow_{k,t-1}^{CN}$ is measured as a strict calendar-day lag; observations following weekends or market holidays do not identify the lagged-flow coefficient. If capital flows systematically anticipate changes in war-related diplomacy, the coefficients on $Flow_{k,t}^{CN}$ and/or $Flow_{k,t-1}^{CN}$ should be statistically significant.

Panel A of Table 8 reports the reverse-causality results. Across columns (1)–(3), neither contemporaneous Stock Connect flows nor strictly calendar-lagged flows predict next-day war-related sentiment. We therefore interpret Panel A as evidence against a direct feedback channel from trading activity to the tone of official Chinese diplomatic rhetoric.

[Table 8 about here.]

We next ask whether the baseline relation reflects low-frequency co-movement between war-related diplomacy and capital flows or higher-frequency variation in rhetoric. To do so, we construct a detrended sentiment index $WDSI_t^{CN}$ by regressing the seven-day raw series on a linear time trend and retaining the residuals, and then re-estimate Equation (1) with this detrended measure. Panel B of Table 8 reports the results. For inflows, the coefficient on detrended sentiment remains close to zero and statistically insignificant. For outflows and net flows, however, the estimates remain economically and statistically strong: the outflow coefficient is negative and significant at the 1% level, and the net-flow coefficient is positive and significant at the 5% level. The detrended specification therefore preserves the same southbound channel as the baseline, indicating that a linear time trend in Chinese war-related diplomatic sentiment does not drive the main result.

C.3. Lag Specification Tests

We next allow for richer dynamics in both capital flows and macro-financial controls. If flows and war-related diplomatic sentiment respond to persistent underlying forces, omitting lagged dependent variables and lagged controls could bias the contemporaneous coefficient on $WDSI_t^{CN}$. We therefore estimate augmented versions of Equation (1) that include one or two lags of the dependent variable together with lagged controls:

$$Flow_{k,t}^{CN} = \alpha_k + \beta_k WDSI_t^{CN} + \phi_{1,k} Flow_{k,t-1}^{CN} + \phi_{2,k} Flow_{k,t-2}^{CN} + \boldsymbol{\gamma}_{0,k}' \mathbf{R}_t + \boldsymbol{\gamma}_{1,k}' \mathbf{R}_{t-1} \\ + \boldsymbol{\gamma}_{2,k}' \mathbf{R}_{t-2} + \boldsymbol{\delta}_{0,k}' \mathbf{Z}_t + \boldsymbol{\delta}_{1,k}' \mathbf{Z}_{t-1} + \boldsymbol{\delta}_{2,k}' \mathbf{Z}_{t-2} + \mu_{m(t)} + \varepsilon_{k,t}, \tag{3}$$

where the first-order specification sets $\phi_{2,k} = 0$, $\boldsymbol{\gamma}_{2,k} = \mathbf{0}$, and $\boldsymbol{\delta}_{2,k} = \mathbf{0}$ (i.e., omits second lags), while the second-order specification estimates these parameters freely.

Panel A of Table 9 reports the first-order specification, and Panel B reports the second-order specification. Across both panels, the coefficient on $WDSI_t^{CN}$ remains small and insignificant for inflows, negative and statistically significant for outflows, and positive and statistically significant for net flows. Capital flows are also strongly persistent, as reflected in the highly significant lagged flow terms and the higher R^2 values relative to the baseline regressions. The main sentiment-flow relation therefore survives the inclusion of richer dynamics, indicating that the baseline estimates are robust to a broad class of lag structures.

[Table 9 about here.]

C.4. Instrumental Variable Strategy

Finally, we use an instrumental variable (IV) strategy to address remaining omitted-variable bias and measurement error in the war-related sentiment index. The design builds on evidence that political and policy elites exhibit persistent, individually specific communication styles that are only imperfectly related to contemporaneous fundamentals. Hansen et al. (2018) document stable heterogeneity in the language used by Federal Open Market Committee (FOMC) members, while Dietrich et al. (2019) show that the emotional intensity of U.S. congressional speeches varies systematically across legislators. Related work in political science and social psychology likewise links gender and other personal characteristics to systematic differences in rhetorical style, including the use of emotional or negative language (Carli, 1990; Haselmayer et al., 2022; Hargrave and Blumenau, 2022). Taken together, this evidence supports our identifying premise that Ministry of Foreign Affairs spokespeople differ persistently in their average "tone."

Building on this evidence, we exploit cross-spokesperson differences in communication style as a source of exogenous variation in war-related diplomatic sentiment. For each spokesperson, we compute the average value of the Chinese sentiment index across all of that spokesperson's answers, regardless of topic, and use this speaker-specific mean, which we label "Spokesperson fixed sentiment", as a proxy for inherent communication style. We then align this cross-sectional measure with the daily time series by constructing a seven-day moving average based on the spokespeople who appear on each day. Intuitively, days on which a systematically more negative (or more positive) spokesperson happens to brief the press should be associated with predictably lower (higher) values of $WDSI_t^{CN}$, even conditional on the same underlying news environment.

The exclusion restriction is that stylistic differences across spokespeople affect Stock Connect flows through their impact on war-related diplomatic sentiment, with no direct effect on fundamentals. This restriction is plausible in our setting. Individual spokespeople have no formal authority over geopolitical decisions, macroeconomic policy, or capital-account regulation, and their assignment to particular briefings is largely administrative. Conditional on the content of the underlying events, which is partly absorbed by the control set and global risk proxies, it is difficult to identify a channel through which investors would respond directly to who delivers the message after controlling for how war-related rhetoric is framed[8]. Under this

[8]This identification logic is closely related to the use of speaker-specific language measures in the FOMC and

restriction, the spokesperson fixed-sentiment measure is a valid instrument for $WDSI_t^{CN}$ in the Stock Connect regressions.

Table 10 reports the IV results. Column (1) presents the first-stage regression of $WDSI_t^{CN}$ on the spokesperson fixed-sentiment instrument and the full set of controls. The coefficient on the instrument is 0.919 with a robust standard error of 0.067, implying a t-statistic of about 13.65 and confirming strong relevance. Across the three second-stage samples, the excluded-instrument first-stage Wald χ^2 statistics range from 182.255 to 186.341. The Kleibergen–Paap rk Wald F statistics range from 184.105 to 188.203, and the rk LM statistics continue to reject under-identification at p-value < 0.001 in all cases. Columns (2)–(4) report the second-stage 2SLS estimates for inflows, outflows, and net flows. The coefficient on $WDSI_t^{CN}$ remains positive but insignificant for inflows, negative and strongly significant for outflows, and positive and strongly significant for net flows. Relative to the baseline OLS estimates, the IV coefficients preserve the same sign but are materially larger in magnitude, especially in the outflow and net-flow regressions, consistent with attenuation bias from classical measurement error in $WDSI_t^{CN}$ pushing the OLS estimates towards zero.

[Table 10 about here.]

Taken together, the placebo, detrending, dynamic, and IV exercises point to the same conclusion. Our baseline findings are unlikely to be driven by generic diplomatic tone, direct contemporaneous reverse causality, low-frequency linear trends, or omitted-variable bias in the war-related sentiment measure. The evidence is therefore consistent with a causal interpretation in which more negative Chinese war-related diplomatic rhetoric predicts stronger southbound

congressional-speech literatures, where individual communication styles are treated as a source of exogenous variation in the tone of public communication (Hansen et al., 2018; Hargrave and Blumenau, 2022).

outflows and weaker net capital inflows.

IV. Additional Analysis

A. Evidence from 14 Foreign Countries

We next ask whether war-related diplomatic sentiment in 14 foreign countries predicts Chinese Stock Connect flows. This cross-country spillover perspective is related to evidence that foreign investors' attention to emerging markets varies with distraction from risk factors in their home markets (Marmora, 2023). Using the daily specification in Equation (1), we regress Chinese northbound inflow, southbound outflow, and net flow on the seven-day moving average of each foreign sentiment index, while holding fixed the same control block and month effects as in the baseline.

Table 11 reports the full-sample estimates. The most stable pattern appears on the southbound margin. Foreign sentiment loads negatively and significantly in the outflow equation for the U.S., U.K., Japan, Brazil, Germany, France, Italy, Mexico, and Russia. Positive net-flow loadings are concentrated in the U.K., Japan, South Korea, Brazil, Germany, Mexico, and Russia, and South Korea and Brazil also enter positively and significantly in the inflow equation. Canada is no longer statistically significant in these daily spillover equations after the repaired Canadian index is used. The evidence therefore points to reallocation through the southbound channel, although the pattern is not universal: Australia loads negatively on inflow, India loads positively on outflow and negatively on net flow, and Spain is largely insignificant. The Indian pattern is consistent with competitive international portfolio allocation between two large emerging markets. Because China and India may be viewed by global investors as offering partly comparable investable growth potential, adverse war-related rhetoric in one market can tilt capital demand towards the other and beyond traditional safe-haven destinations. These spillover estimates complement the baseline China result in Section III.A and the relative-price evidence based on the A-H share premium in Section IV.B, both of which likewise indicate reallocation through the southbound channel when war-related risk intensifies.

[Table 11 about here.]

We next ask whether a country's own war-related diplomatic tone predicts its own monthly international equity position. Table 12 uses monthly means of the full-signed daily WDSI series: Panel A is based on the seven-day WDSI, and Panel B is based on the thirty-day WDSI. The specification is deliberately narrow. It keeps both positive and negative daily WDSI values and excludes raw daily scores, monthly low-points, and negative-part transformations. The United States supplies the clear safe-haven evidence. Its WDSI coefficient is negative in the inflow and net-flow columns in both panels; because lower WDSI denotes more adverse rhetoric, these estimates imply larger inflows and a stronger net equity position when U.S. diplomatic tone worsens. South Korea points in the opposite direction: its positive coefficients imply weaker inflows and weaker net positions when Korean WDSI falls. Estimates for Japan, Brazil, and Canada are statistically weak in this specification. The proxy rows reinforce the limited scope of the result. France's 12-month portfolio-balance proxy is indistinguishable from zero, Italy shows a larger outflow response only in the seven-day panel, and Russia's OFZ-demand proxy falls when sentiment worsens, which is inconsistent with safe-haven demand. Taken together, the monthly own-country evidence supports a U.S.-centered safe-haven channel and offers little evidence of a broad tendency for adverse domestic rhetoric to attract equity capital.

[Table 12 about here.]

Figure 6 translates the monthly pairwise regressions into a map of significant source-target links. Each arrow is a separate predictive relation from the source country's WDSI to the target country's equity-flow outcome, estimated with the same monthly control structure; the figure should therefore be read as a set of bilateral channels. The gross-flow panels place South Korea near the center of the network. Chinese, Japanese, and U.S. sentiment predict Korean inflows, and the same three source countries also predict Korean outflows; Korean sentiment predicts U.S. and Brazilian inflows. The net-flow panel shifts the receiving margin towards the United States. Chinese, Japanese, and Canadian sentiment load negatively on U.S. net flows, while Brazilian sentiment loads positively. Brazil and Canada also appear as source or target nodes in several bilateral links, even though their own-country coefficients in Table 12 are statistically weak. The message is one of selective transmission: monthly WDSI shocks travel through a small number of directed country pairs, with South Korea prominent in gross flows and the United States prominent in net flows.

[Figure 6 about here.]

B. Relative-Price Evidence: The A-H Share Premium

Following He et al. (2023), we examine one relative-price channel through which war-related diplomatic sentiment may affect cross-border equity flows: the A-H share premium. When geopolitical and war-related risks intensify, international investors may reduce their exposure to Hong Kong-listed Chinese firms, putting downward pressure on H-share prices. If mainland A-shares remain relatively supported by domestic investors, the A-H premium widens and creates an arbitrage incentive for mainland investors to buy cheaper H-shares through the southbound Stock Connect channel. This relative-price margin helps explain why the Chinese war-related sentiment effect is strongest in southbound reallocation. It also provides a way to infer the response of international investors in the Hong Kong market even though their trades are not directly observed.

We test this relative-price channel by regressing the daily A-H premium on the Chinese war-related diplomatic sentiment index $WDSI^{CN}$. The specification controls for Chinese, U.S., and developed-market equity returns, Chinese and U.S. geopolitical risk, Chinese and U.S. economic policy uncertainty, gold returns, the RMB-USD exchange rate, and month-of-year fixed effects. Column (1) of Table 13 shows that the coefficient on $WDSI^{CN}$ is negative and highly significant: a one-unit increase in $WDSI^{CN}$ lowers the A-H premium by about 0.89 percentage points. Because lower values of $WDSI^{CN}$ correspond to more adverse war-related rhetoric, the estimate implies that heightened war risk raises the A-H premium and makes H-shares cheaper relative to their A-share counterparts.

[Table 13 about here.]

The A-H premium is highly persistent, so columns (2) and (3) add one and two lags of the premium to the specification. In column (3), the two lag coefficients sum to 0.994, indicating near-unit-root persistence in the level series. Untabulated ADF and KPSS tests are consistent with nonstationary level behaviour for the A-H premium. The dynamic specification nevertheless addresses this concern. The residual from column (3) strongly rejects a unit root in the ADF test and does not reject stationarity in the KPSS test at the 5% level, and the coefficient on $WDSI^{CN}$ remains negative and statistically significant at the 10% level under Newey–West HAC inference. Column (4) shows that the result also survives in first differences: when the dependent variable is the daily change in the A-H premium, the coefficient on $WDSI^{CN}$ is -0.075 and remains significant at the 10% level. These checks reduce the concern that nonstationary residuals generate the finding. Negative shocks to Chinese war-related diplomatic sentiment move the A-H premium over and above its own past history, consistent with war-related rhetoric shifting relative prices between mainland and Hong Kong shares in a way that opens arbitrage opportunities for mainland investors.

This relative-price evidence also reduces concerns about reverse causality from capital flows to the A-H premium. If capital flows were the primary driver of the premium, heightened war-related risk would more naturally predict capital outflows from the mainland and an A-share discount. The estimates point in the opposite direction: more adverse war-related sentiment is associated with a higher A-H premium. The more plausible interpretation is that diplomatic sentiment first moves the relative price between A- and H-shares, which then encourages additional southbound investment into cheaper H-shares and links the baseline flow results to a concrete relative-price channel.

We further assess directionality with supplementary daily Granger-causality tests between the A-H premium and Stock Connect flows in a residualised VAR that uses the same control block and month-of-year fixed effects as the baseline. The lag-1/3/5 results in Table IV.3 (Online Appendix IV) show that the A-H premium significantly predicts southbound outflow, whereas the reverse-direction evidence for outflow is weak. This pattern supports the interpretation that the premium acts as a transmission channel for flow reallocation, beyond serving as a contemporaneous correlate of Stock Connect trading.

The A-H premium is the relative-valuation wedge most directly tied to southbound reallocation. We next examine whether the same Chinese war-related sentiment shocks are also reflected in broader equity-index returns.

C. Complementary Return Evidence: Beyond an Absolute-Price-Decline Channel

The main empirical analysis studies cross-border equity flows. We use index returns as complementary price-side evidence on whether the same Chinese war-related sentiment shifts are reflected in broader equity-market repricing. We estimate four return specifications with the control block in Equation (1): (i) a daily specification using lagged baseline sentiment, $WDSI_{t-1}^{CN}$, constructed from the seven-day moving average with a one-day lag; (ii) a weekly specification using baseline sentiment, $WDSI_{t-1}^{CN}$, constructed from the seven-day moving average; (iii) a weekly specification using the alternative thirty-day sentiment measure, $WDSI_{t-1}^{CN,30}$; and (iv) a monthly specification using contemporaneous thirty-day sentiment, $WDSI_{t}^{CN,30}$.

Table 14 provides little support for a simple absolute-price channel in which the flow response is driven by investors buying after broad market declines. Panel A shows that the daily return coefficients are uniformly negative, small, and statistically imprecise. The weekly estimates in Panels B and C remain negative across indices, with stronger precision under the thirty-day sentiment measure, and Panel D shows large and statistically significant negative coefficients at the monthly horizon. Because lower values of $WDSI$ indicate more adverse war-related diplomatic sentiment, these negative coefficients imply that worsening sentiment is associated with higher Chinese equity-index returns. Figure 7 shows the same pattern: daily estimates cluster near zero, weekly estimates shift to the negative side, and monthly estimates are farthest from zero. Together with the baseline flow results, this evidence points to portfolio reallocation and relative-pricing adjustment, with limited support for a mechanical reaction to broad equity-market declines.

[Table 14 about here.]

[Figure 7 about here.]

D. Market Activity and Risk Absorption

The preceding analysis focuses on the direction and magnitude of cross-border equity flows. War-related diplomatic sentiment may also matter on a different margin: how intensely investors trade and how risk is absorbed within domestic equity markets. Market activity, especially trading volume and return volatility, is therefore a complementary dimension of geopolitical-risk transmission. War-related news can trigger rebalancing, hedging, or speculative trading even when aggregate net flows are modest (Bollerslev et al., 2018). We test this channel by relating daily Chinese and U.S. war-related diplomatic sentiment to trading volume and realized volatility in the corresponding equity markets.

Panel A of Table 15 reports regressions of log daily trading volume on war-related diplomatic sentiment. Columns (1)–(2) show negative and highly significant coefficients on the Chinese sentiment index $WDSI^{CN}$ for both the CSI 300 and CSI 1000. Because lower values of $WDSI^{CN}$ correspond to more negative Chinese war-related rhetoric, a one-unit decrease in $WDSI^{CN}$ is associated with approximately a 5.3% increase in CSI 300 trading volume and a 10.7% increase in CSI 1000 trading volume. Columns (3)–(4) show the same pattern for U.S. markets: a one-unit decrease in $WDSI^{US}$ is associated with a 14.4% increase in Nasdaq trading volume and a 11.2% increase in Dow Jones Industrial Average trading volume. Across both markets, adverse war-related rhetoric is associated with more active trading, consistent with portfolio rebalancing and risk-management demand.

[Table 15 about here.]

Panel B of Table 15 turns to realized volatility. In the U.S. columns (7)–(8), the coefficients on $WDSI^{US}$ are negative and highly significant, so a deterioration in U.S. war-related rhetoric predicts higher volatility on both Nasdaq and the Dow Jones Industrial Average. The Chinese volatility regressions display a different risk-absorption pattern. The coefficients on $WDSI^{CN}$ in columns (5)–(6) are positive and statistically significant, implying that more adverse Chinese war-related rhetoric is associated with lower realized volatility for the CSI 300 and CSI 1000, even as trading volumes rise. The volume response is common across markets, while the volatility response differs: adverse rhetoric raises both turnover and volatility in U.S. markets, whereas in onshore China it raises turnover while dampening measured daily volatility. This pattern is consistent with a combination of risk-absorbing flows and policy-driven volatility smoothing, and shows that war-related diplomatic sentiment affects cross-border capital flows, domestic trading, and volatility margins.

V. Conclusions

This paper develops high-frequency war-related diplomatic sentiment indices from official foreign-policy communications for the 15 largest world economies. Relative to newspaper-based geopolitical risk measures, the WDSI provides higher-frequency official-policy information and displays stronger short-to-medium-horizon predictive ability for future China-specific GPR than the reverse GPR-to-WDSI direction. The empirical evidence focuses on China because its Stock Connect system provides uniquely informative daily, directionally decomposed international equity-flow data. More negative Chinese war-related diplomatic sentiment predicts stronger southbound reallocation and weaker net Stock Connect flows. Foreign WDSIs also transmit primarily through China's southbound leg in daily spillover regressions. Among economies with sufficiently informative monthly own-country capital-flow data, the United States is the only core market with a clear safe-haven pattern; South Korea moves away from a safe-haven interpretation, while estimates for Canada, Japan, and Brazil are largely insignificant. The repaired Canadian result provides weak inflow evidence: its positive inflow coefficients are statistically imprecise, and lower Canadian WDSI has no robust association with own-country inflows. The six-country pairwise network shows that cross-country transmission is directional and concentrated, with the United States and South Korea as the main receiving nodes and Brazil and Canada contributing economically meaningful links to the core structure. Complementary price-side tests further argue against a simple absolute-price-decline channel: daily Chinese return responses are weak, whereas the A-H premium, market activity, and medium-horizon repricing all remain responsive to war-related sentiment.

These findings matter for both investors and policymakers. For investors, official diplomatic language contains forward-looking information about cross-border reallocation, and the cross-country evidence cautions against treating markets as symmetric safe havens. The United States absorbs adverse diplomatic shocks differently from China, South Korea, and other non-U.S. markets. More broadly, the evidence suggests that only markets with the strongest safe-haven status can stably attract capital when their own war-related diplomatic rhetoric turns more negative; other markets tend to experience capital outflows, reduced inflows, or statistically weak responses, with no evidence of a comparable inflow benefit. This redistribution channel offers a capital-flow perspective on the economic motives that may surround geopolitical risk, because adverse official rhetoric changes perceived risk and the international allocation of portfolio demand. For policymakers, official communication is a market-relevant signal whose effects propagate through both domestic and foreign portfolio channels. The paper shows that official diplomatic rhetoric can be transformed into a transparent empirical state variable for studying international portfolio reallocation, relative-price adjustment, and market activity in real time.

References

Bailey, M. A., A. Strezhnev, and E. Voeten (2017). Estimating dynamic state preferences from united nations voting data. *International Organization 71*(1), 1–31.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics 131*(4), 1593–1636.

Baturo, A., N. Dasandi, and S. J. Mikhaylov (2017). Understanding state preferences with text as data: Introducing the UN General Debate corpus. *Research & Politics 4*(2), 1–9.

Baum, M. A. and Y. M. Zhukov (2019). Media ownership and news coverage of international conflict. *Political Communication 36*(1), 1–22.

Baur, D. G. and T. K. McDermott (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance 34*(8), 1886–1898.

Benhima, K. and R. Cordonier (2022). News, sentiment and capital flows. *Journal of International Economics 137*, 103621.

Bohn, H. and L. L. Tesar (1996). U.s. equity investment in foreign markets: Portfolio rebalancing or return chasing? *American Economic Review 86*(2), 77–81.

Bollerslev, T., J. Li, and Y. Xue (2018). Volume, volatility, and public news announcements. *Review of Economic Studies 85*(4), 2005–2041.

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

Buehlmaier, C. and T. M. Whited (2018). Are financial constraints priced? evidence from textual analysis. *Review of Financial Studies 31*(7), 2693–2728.

Caldara, D. and M. Iacoviello (2022). Measuring geopolitical risk. *American Economic Review 112*(4), 1194–1225.

Camanho, N., H. Hau, and H. Rey (2022). Global portfolio rebalancing and exchange rates. *Review of Financial Studies 35*(11), 5228–5274.

Cao, C., X. Li, and G. Liu (2019). Political uncertainty and cross-border acquisitions. *Review of Finance 23*(2), 439–470.

Caporale, G. M., F. Menla Ali, and N. Spagnolo (2015). Exchange rate uncertainty and international portfolio flows: A multivariate garch-in-mean approach. *Journal of International Money and Finance 54*, 70–92.

Carli, L. L. (1990). Gender, language, and influence. *Journal of Personality and Social Psychology 59*(5), 941–951.

Choi, S. and J. Havel (2025). Geopolitical risk and U.S. foreign portfolio investment: A tale of advanced and emerging markets. *Journal of International Money and Finance 151*, 103253.

Correa, R., K. Garud, J. M. Londono, and N. Mislang (2021). Sentiment in central banks' financial stability reports. *Review of Finance 25*(1), 85–120.

Dietrich, B. J., T. J. Hayes, and D. Z. O'Brien (2019). Pitch perfect: Vocal pitch and the emotional intensity of congressional speech. *American Political Science Review 113*(4), 941–962.

Feng, C., L. Han, S. A. Vigne, and Y. Xu (2023). Geopolitical risk and the dynamics of international capital flows. *Journal of International Financial Markets, Institutions & Money 82*, 101693.

Froot, K. A., P. G. J. O'Connell, and M. S. Seasholes (2001). The portfolio flows of international investors. *Journal of Financial Economics 59*(2), 151–193.

Fuchs, A. and N.-H. Klann (2013). Paying a visit: The Dalai Lama effect on international trade. *Journal of Development Economics 101*, 55–63.

Gartzke, E. (2006). The affinity of nations: Similarity of state voting positions in the UN general assembly, 1946–2005. Dataset, Version 4.0. University of California, San Diego.

Gelman, M., A. Jochem, S. Reitz, and M. P. Taylor (2015). Real financial market exchange rates and capital flows. *Journal of International Money and Finance 54*, 50–69.

Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as data. *Journal of Economic Literature 57*(3), 535–574.

Gowa, J. (1994). *Allies, Adversaries, and International Trade*. Princeton, NJ: Princeton University Press.

Griffin, J. M., F. Nardari, and R. M. Stulz (2004). Are daily cross-border equity flows pushed or pulled? *Review of Economics and Statistics 86*(3), 641–657.

Hansen, S., M. McMahon, and A. Prat (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *Quarterly Journal of Economics 133*(2), 801–870.

Hargrave, E. and J. Blumenau (2022). No longer conforming to stereotypes? gender, political style and parliamentary debate in the UK. *British Journal of Political Science 52*(4), 1584–1603.

Haselmayer, M., S. C. Dingler, and M. Jenny (2022). How women shape negativity in parliamentary speeches—a sentiment analysis of debates in the Austrian parliament. *Parliamentary Affairs 75*(4), 867–886.

Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun (2019). Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics 134*(4), 2135–2202.

Hau, H. and H. Rey (2006). Exchange rates, equity prices, and capital flows. *Review of Financial Studies 19*(1), 273–317.

He, Z., Y. Wang, and X. Zhu (2023). The Stock Connect to China. In *AEA Papers and Proceedings*, Volume 113, Nashville, TN, pp. 125–130. American Economic Association.

Hufbauer, G. C., J. J. Schott, K. A. Elliott, and B. Oegg (2007). *Economic Sanctions Reconsidered* (3 ed.). Washington, DC: Peterson Institute for International Economics.

Julio, B. and Y. Yook (2012). Political uncertainty and corporate investment cycles. *Journal of Finance 67*(1), 45–83.

Kelly, B., L. Pastor, and P. Veronesi (2016). The price of political uncertainty: Theory and evidence from the option market. *Journal of Finance 71*(5), 2417–2480.

Kommandeur, J. and R. van Dam (2024, July). Gina diplomatic: Methodological notes. Technical report, The Hague Centre for Strategic Studies (HCSS). HCSS Datalab. Created/updated July 2024.

Lundblad, C. T., D. Shi, X. Zhang, and Z. Zhang (2026). Foreign capital in the Chinese stock market: A firm-level study. *Journal of Financial and Quantitative Analysis 61*(2), 799–840.

Maoz, Z. (2010). *Networks of Nations: The Evolution, Structure, and Impact of International Networks, 1816–2001.* Cambridge: Cambridge University Press.

Marmora, P. (2023). Identifying the effects of macroeconomic attention through foreign investor distraction. *Journal of Financial and Quantitative Analysis 58*(8), 3644–3671.

Mochtak, M. and R. Q. Turcsanyi (2021). Studying chinese foreign policy narratives: Introducing the ministry of foreign affairs press conferences corpus. *Journal of Chinese Political Science 26*(4), 743–761.

Papaioannou, E. (2009). What drives international bank flows? politics, institutions and other determinants. *Journal of Development Economics 88*(2), 269–281.

Pástor, L. and P. Veronesi (2012). Uncertainty about government policy and stock prices. *Journal of Finance 67*(4), 1219–1264.

Permutable AI (2024). War index: What wars are happening now 2025 update. Web page (shows original date Jul 22, 2024; later updates shown on page). Accessed 2025-12-29.

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

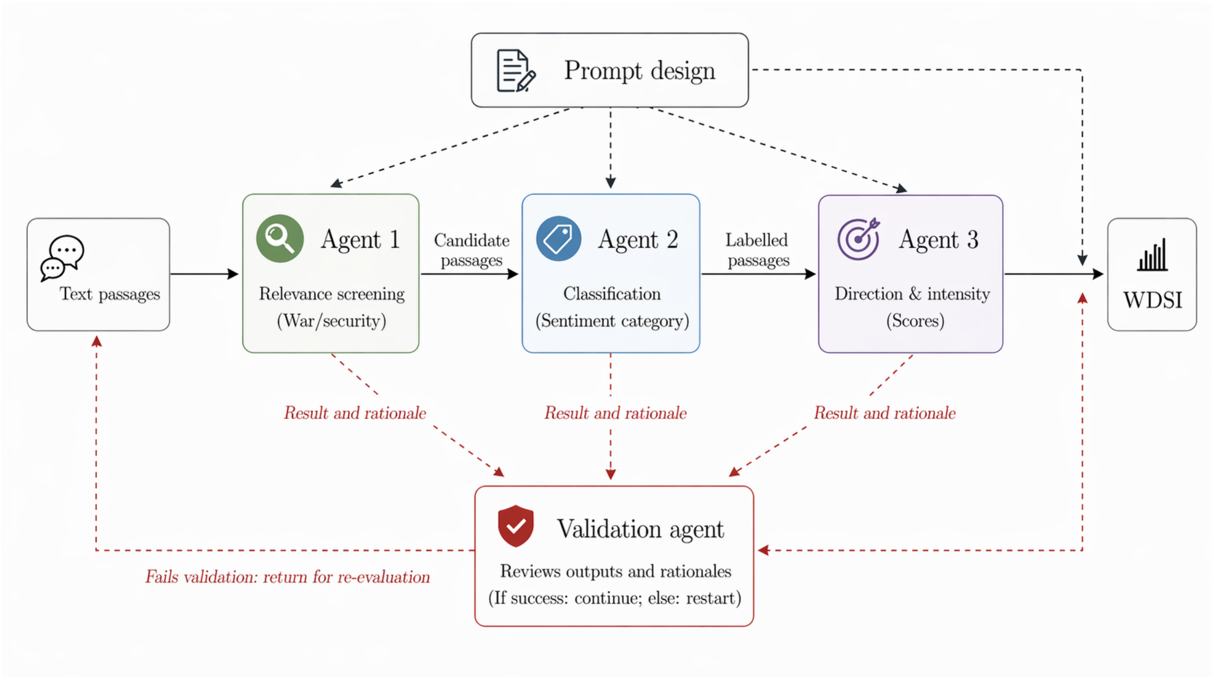

Figure 1: Process of constructing the War-Related Diplomatic Sentiment Index.

Note: The workflow consists of three sequential LLM stages (war-related screening, sentiment classification, and intensity scoring), followed by a validation agent. The input is a source-specific text passage: a Q&A exchange for Chinese MFA press conferences and comparable briefings, or a self-contained release, statement, remark, readout, speech, or source-provided section for non-Q&A corpora. Solid arrows indicate the primary processing sequence, while dashed arrows transmit intermediate outputs and rationales to the validator. If validation fails, the item is routed back to the input stage for re-evaluation.

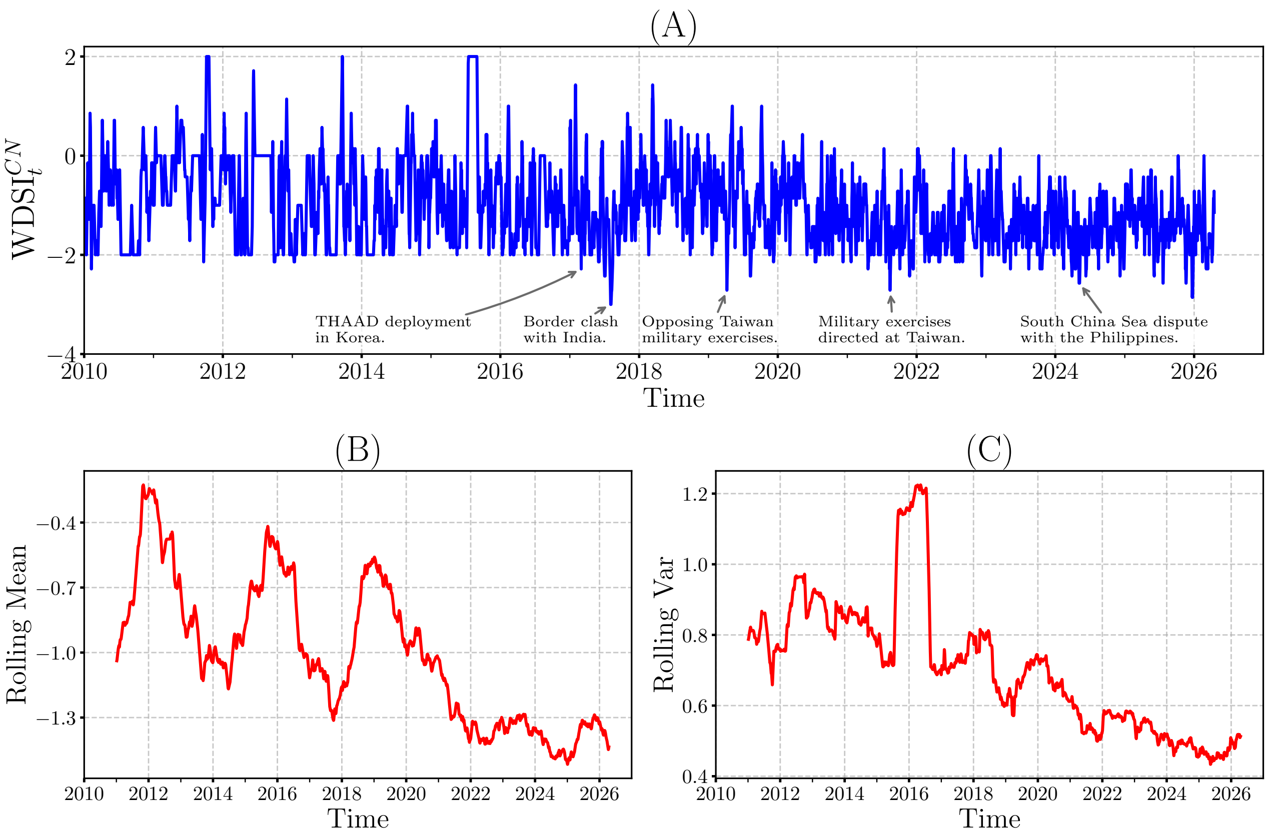

Figure 2: Chinese War-Related Diplomatic Sentiment Index ($WDSI^{CN}$): Daily Series, Rolling Mean, and Rolling Volatility

Panel A plots the daily Chinese WDSI over the sample period (7 January 2010 to 17 April 2026), constructed as a seven-day rolling average of the raw per-conference sentiment scores on the [-3, 3] scale. Several major events in the military and security domain (for example U.S. deployment of missile defense in Korea, border clashes with India, large-scale military exercises around Taiwan and maritime disputes in the South China Sea) are annotated along the timeline. Panel B reports the one-year rolling mean of the WDSI, while Panel C shows the corresponding one-year rolling variance. The decline in volatility and the gradual downward drift of the rolling mean indicate that China's official diplomatic language has become progressively more negative on military and security issues over time.

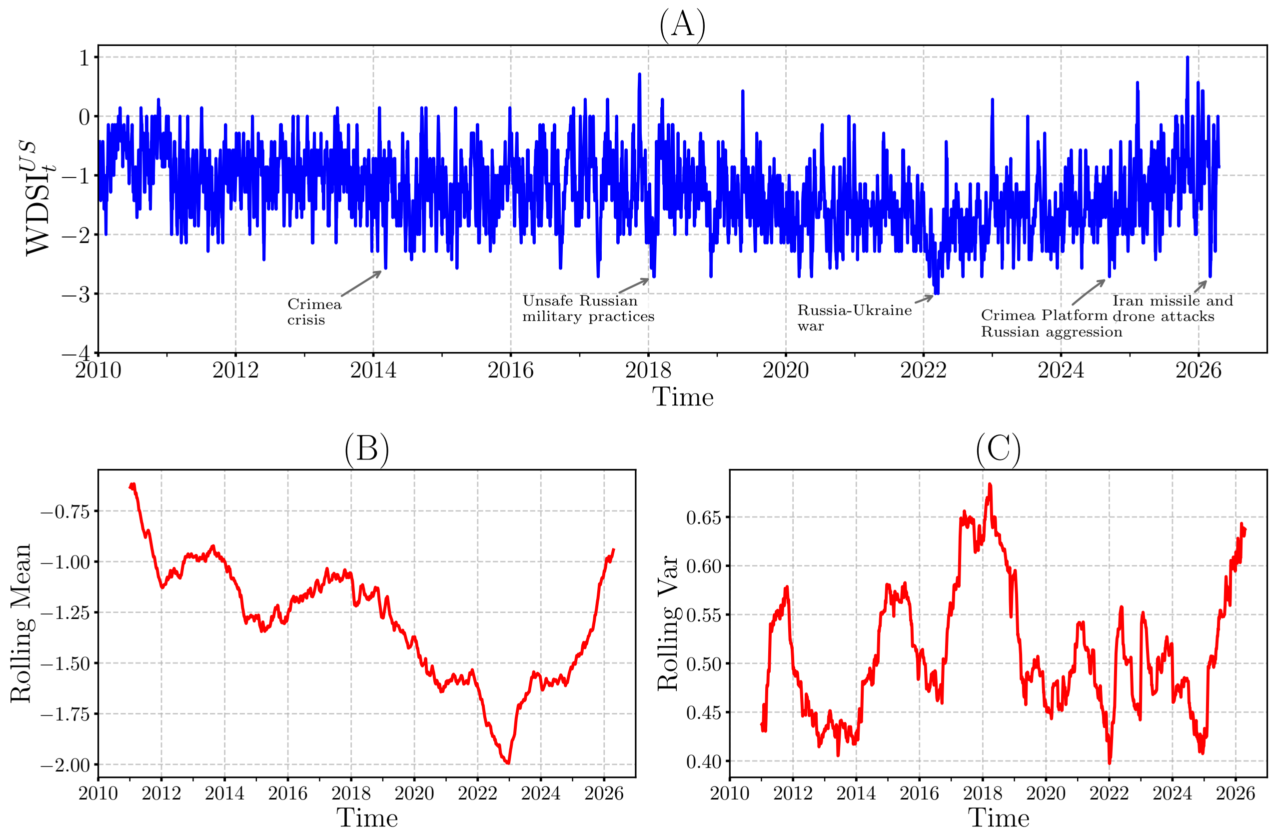

Figure 3: U.S. War-Related Diplomatic Sentiment Index ($WDSI^{US}$): Daily Series, Rolling Mean, and Rolling Volatility

Panel A plots the daily U.S. WDSI over the sample period (7 January 2010 to 18 April 2026), constructed as a seven-day rolling average of the raw sentiment scores on the [-3, 3] scale derived from Department of State official releases. Several low-point security episodes (the Crimea crisis, unsafe Russian military practices, the Russia-Ukraine war, Crimea Platform statements on Russian aggression, and Iran missile and drone attacks) are annotated along the timeline. Panel B reports the one-year rolling mean of the U.S. WDSI, which fluctuates around a moderately negative level and becomes more negative during episodes of heightened military involvement and geopolitical tension. Panel C shows the corresponding one-year rolling variance. Volatility is clearly time-varying, with pronounced spikes around periods of sustained conflict and crisis, indicating that the intensity of U.S. war-related diplomatic sentiment is clustered in time and unevenly distributed over the sample.

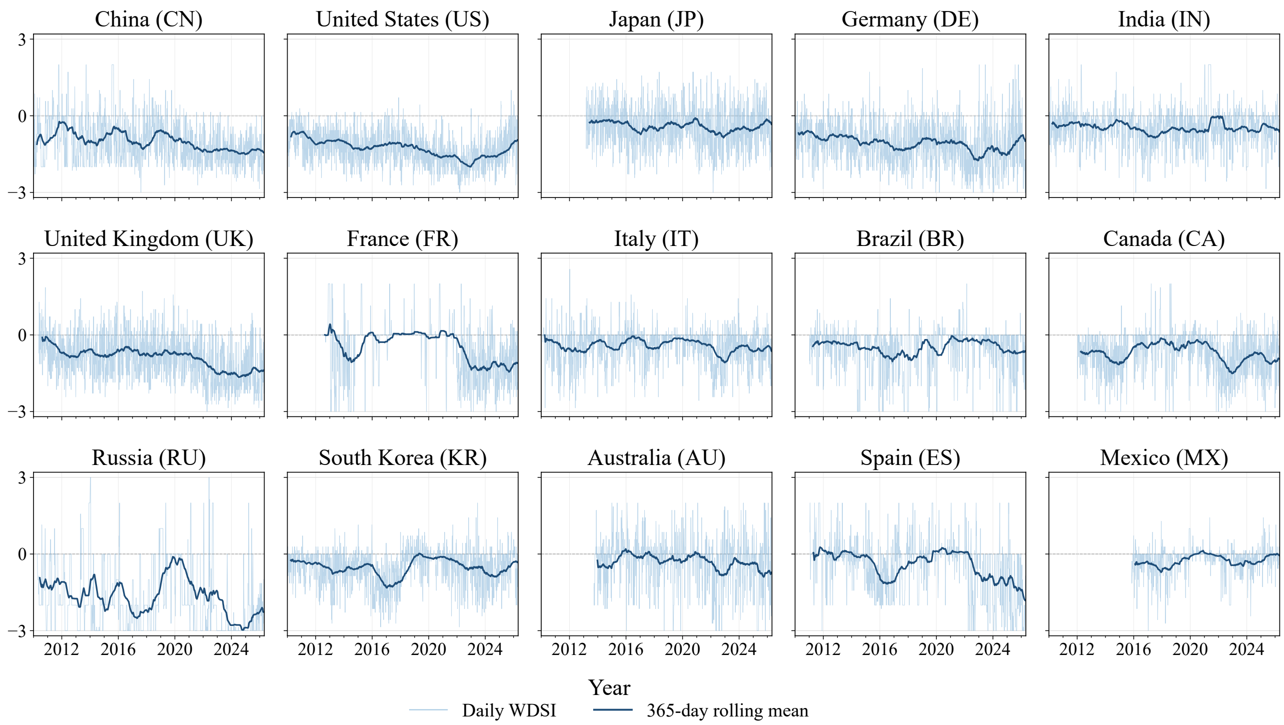

Figure 4: War-Related Diplomatic Sentiment across 15 Economies

This figure plots the daily seven-day moving-average WDSI and its 365-day rolling mean for the 15 largest world economies in the official diplomatic-text database. Lower values indicate more negative war-related diplomatic sentiment. Panels enter when the corresponding country corpus becomes available, so early blank regions reflect source-coverage windows and should not be read as zero sentiment. For France, the segmented appearance reflects the publication calendar and archive coverage of the France Diplomatie source, including intervals with few usable official texts or unavailable page bodies in the local corpus; these intervals are source-coverage features and should not be read as zero sentiment observations. The common vertical scale preserves cross-country comparability, while the rolling mean highlights persistent shifts in official diplomatic tone.

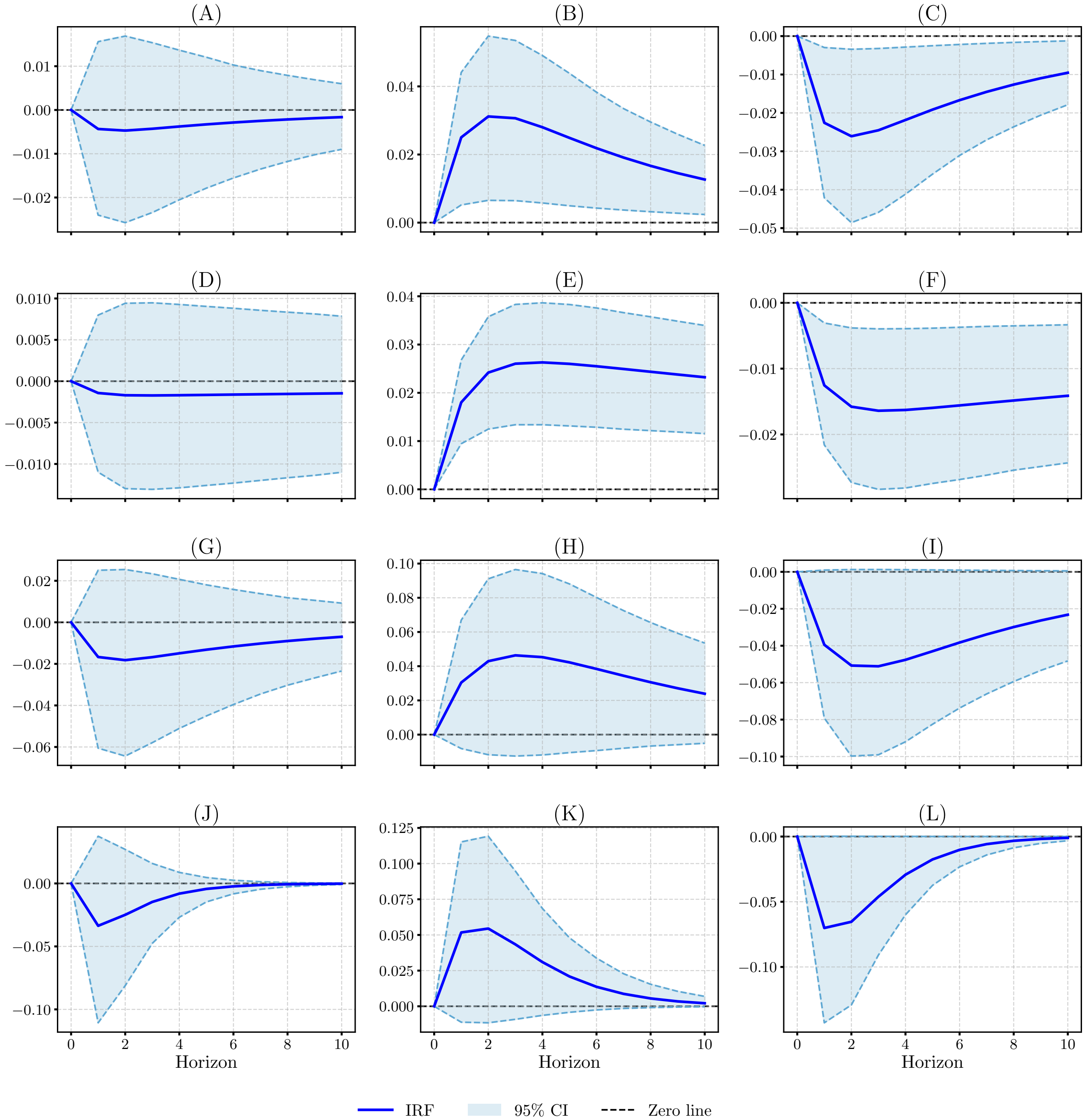

Figure 5: Impulse Responses to a Negative Sentiment Shock: VAR(1), Daily and Weekly Specifications

The figure reports orthogonalised impulse responses for inflow, outflow, and net flow under four residualised VAR(1) specifications. Panels A-C use daily baseline $WDSI_t^{CN}$ (seven-day moving average). Panels D-F use daily $WDSI_t^{CN,30}$ (thirty-day moving average). Panels G-I use weekly $WDSI_t^{CN,30}$, and Panels J-L use weekly $WDSI_t^{CN}$. Weekly observations are built from Monday-anchored weeks, with flow variables summed and sentiment/control variables averaged within week. Shaded areas and dashed bounds indicate 95% confidence intervals, and the dashed horizontal line marks zero. Responses are for a one-standard-deviation negative sentiment shock (a decline in the sentiment index).

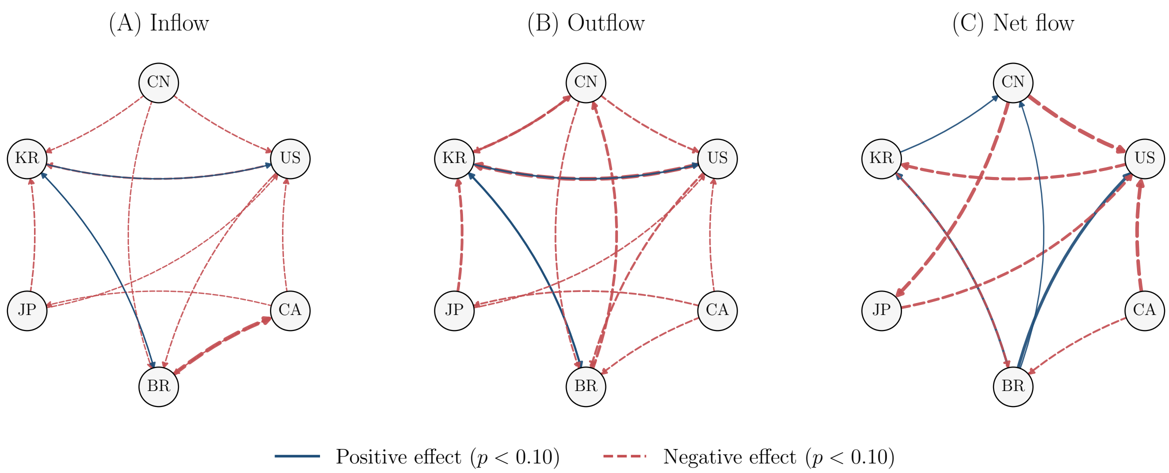

Figure 6: Directional Cross-Country Sentiment-Flow Network (Core 6: CN, US, JP, KR, BR, CA)

This figure shows directional cross-country links from war-related diplomatic sentiment to monthly equity-flow outcomes among the six-country core set: China (CN), the United States (US), Japan (JP), South Korea (KR), Brazil (BR), and Canada (CA). Panels A-C report results for monthly Inflow, Outflow, and Net flow, respectively. For each ordered source-target pair and each target-country flow outcome, we estimate a separate monthly regression in which the dependent variable is the target-country flow measure and the key explanatory variable is the source country's monthly mean of the full-signed daily seven-day moving-average WDSI series; all specifications include the same common monthly control block (U.S. and developed-market returns, gold returns, global U.S.-based GPR and EPU measures) and month-of-year fixed effects, with heteroskedasticity-robust (HC1) standard errors. A directed arrow from country j to country i is plotted only when the corresponding sentiment coefficient is significant at the 10% level (p-value < 0.10); self-links are omitted. Blue solid arrows denote positive coefficients and red dashed arrows denote negative coefficients. Arrow thickness scales with the absolute coefficient magnitude within each panel. Node positions and curve shapes are chosen for visual clarity only and do not carry econometric interpretation.

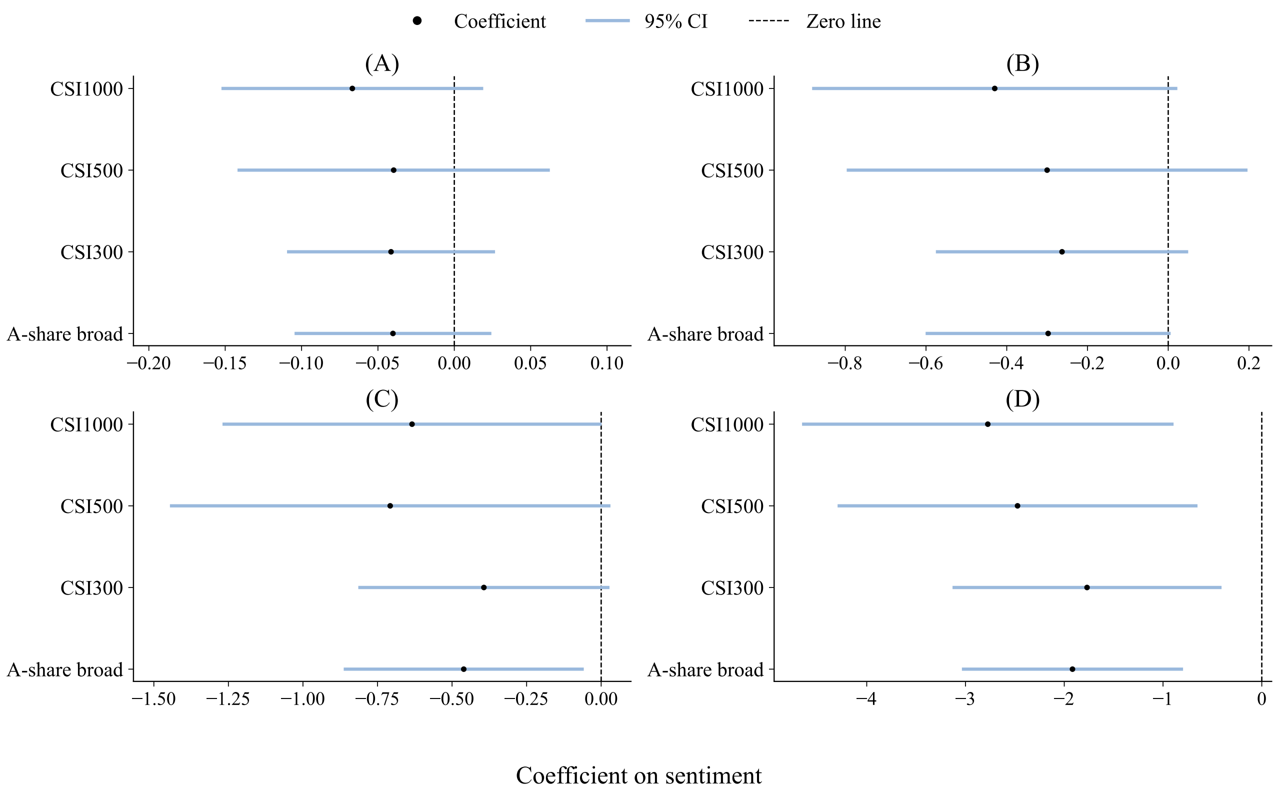

Figure 7: Sentiment Coefficients for Chinese Equity Returns

The figure plots sentiment coefficients and 95% confidence intervals from Table 14. Panel A reports daily baseline $WDSI_{t-1}^{CN}$ (7-day MA, one-day lag). Panel B reports weekly baseline $WDSI_{t-1}^{CN}$ (7-day MA). Panel C reports weekly alternative $WDSI_{t-1}^{CN,30}$ (30-day MA). Panel D reports monthly contemporaneous $WDSI_{t}^{CN,30}$.

Table 1: Distribution of Chinese WDSI Scores and Example of Model Output

Panel A reports the distribution of Chinese WDSI scores assigned to individual press-conference days on the [-3, 3] scale, before any forward-filling or temporal interpolation. Panels B and C present illustrative negative (-3) and positive (3) cases, respectively, including the input text, date and press-conference title, assigned score, and the AI-generated rationale.

Panel A. Distribution of Raw Discrete Chinese WDSI Scores

	-3	-2	-1	0	1	2	3
Number of press-conference days	76	1,533	635	552	184	112	3
Share (%)	2.46	49.53	20.52	17.84	5.95	3.62	0.10

Panel B. Example of Negative War-related Diplomatic Sentiment (Score: -3; Date: 2022-12-09)

Title	Regular Press Conference of Foreign Ministry Spokesperson Mao Ning, 9 December 2022
Q&A Unit	**Agence France-Presse Reporter:** The U.S. House of Representatives has passed the "National Defense Authorization Act", which includes provisions for up to US$10 billion in military loans to Taiwan over the next five years. What is your comment? **Spokesperson Ning Mao:** China firmly opposes the inclusion of negative China-related content in the NDAA. Taiwan is China's Taiwan, and the Taiwan question is purely China's internal affair. The U.S. should abide by the one-China principle and the three China-U.S. joint communiques, delete the negative China-related content in the bill, stop playing the "Taiwan card" and using Taiwan to contain China, and stop U.S.-Taiwan military contacts so as to avoid further damage to China-U.S. relations. China will take all necessary measures to firmly safeguard its sovereignty and security interests.
Model Rationale	The exchange condemns the U.S. "National Defense Authorization Act" for authorising large-scale military loans to Taiwan, reaffirms that Taiwan is part of China and that the Taiwan issue is a purely internal matter, urges the U.S. to stop "playing the Taiwan card" and all U.S.-Taiwan military contacts, and warns that China will take "all necessary measures" to defend its sovereignty and security. The combination of strong negative evaluation, opposition to foreign military involvement, and threats of countermeasures leads the pipeline to classify this unit as high-intensity negative war-related sentiment and to assign the lowest score of -3.

Panel C. Example of Positive War-related Diplomatic Sentiment (Score: 3; Date: 2017-01-01)

Title	Response by Foreign Ministry Spokesperson Chunying Hua to a Question on the Istanbul Shooting Incident, 1 January 2017
Q&A Unit	**Reporter:** In the early hours of 1 January local time, a shooting occurred at a nightclub in Istanbul, Turkey, causing at least 39 deaths and more than 60 injuries. The Turkish government said the incident was a terrorist attack. What is China's comment? **Spokesperson Chunying Hua:** Foreign Minister Wang Yi has sent a message of condolences to the Turkish foreign minister over the shooting. China strongly condemns this incident, firmly opposes all forms of terrorism, and is willing to strengthen coordination and cooperation with Turkey and the international community to safeguard regional and world peace and security.
Model Rationale	The unit combines unequivocal condemnation of terrorism with explicit support for coordinated international action to maintain regional and world peace and security. The final scoring file therefore classifies this unit as high-intensity positive war-related diplomatic sentiment and assigns the highest score of 3.

Table 2: Directional VAR Evidence Between Monthly Low-Point WDSI and GPR

This table reports monthly bivariate VAR(p) results for China's low-point war-related diplomatic sentiment index $WDSI_m^{CN,\min}$ and the China-specific geopolitical risk index GPR_m^{CN} over lag orders $p = 1, \ldots, 12$. Lower values of $WDSI_m^{CN,\min}$ indicate more adverse official war-related sentiment in that month. The estimation uses raw (non-residualised) level series over January 2010 to March 2026, with no additional controls included inside the VAR system. "p-value: WDSI→GPR" is the Granger-causality test for the null that lagged WDSI does not predict GPR. "p-value: GPR→WDSI" reports the reverse-direction test. Lower p-values indicate stronger Granger predictability; the WDSI→GPR direction is significant across a broader lag range than the reverse direction. Lower AIC/BIC indicate better in-sample fit. All estimated VAR systems are stable. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Lag order p	Observations	p-value: WDSI→GPR	p-value: GPR→WDSI	AIC	BIC
1	194	0.008***	0.131	-3.780	-3.679
2	193	0.000***	0.006***	-3.858	-3.689
3	192	0.002***	0.030**	-3.826	-3.588
4	191	0.004***	0.146	-3.839	-3.532
5	190	0.010**	0.242	-3.807	-3.431
6	189	0.039**	0.314	-3.770	-3.324
7	188	0.064*	0.276	-3.747	-3.231
8	187	0.092*	0.398	-3.701	-3.113
9	186	0.082*	0.642	-3.666	-3.007
10	185	0.104	0.578	-3.641	-2.910
11	184	0.103	0.668	-3.643	-2.839
12	183	0.184	0.913	-3.640	-2.763

Table 3: Descriptive Statistics

The table reports detailed summary statistics for the variables used in the baseline China regressions and for the seven-day moving-average WDSI series for all 15 countries. The WDSI rows use each country's available daily WDSI sample. Stock Connect flow variables are summarised over their channel-specific nonmissing China flow samples. Control variables are summarised over their available aligned daily control sample after transformation and winsorisation, while the baseline regressions use the dependent-variable-specific listwise intersection of flows, WDSI, controls, and fixed effects. Accordingly, the reported Obs. varies across rows and should not be read as a common regression sample. Flow variables are measured in hundred million USD; equity returns are expressed in percentage points; gold returns are log first differences and are reported in decimal units. Non-WDSI variables are transformed and winsorised consistently with the baseline specifications.

Variable (Full Name)	Symbol	Unit	Obs.	Mean	Std. Dev.	Median	Min.	Max.
Net Northbound Inflows	Inflow	USD 100m	2,205	0.988	4.905	0.768	-9.023	11.496
Net Southbound Flows	Outflow	USD 100m	2,224	1.815	2.834	1.198	-2.827	8.632
Net Stock Connect Flow	NetFlow	USD 100m	2,190	-0.782	6.179	-0.483	-14.574	11.062
China War-Related Diplomatic Sentiment	$WDSI^{CN}$	Index	5,945	-1.019	0.798	-1.143	-3.000	2.000
United States War-Related Diplomatic Sentiment	$WDSI^{US}$	Index	5,946	-1.272	0.602	-1.286	-3.000	1.000
Japan War-Related Diplomatic Sentiment	$WDSI^{JP}$	Index	4,793	-0.423	0.740	-0.429	-3.000	1.714
Germany War-Related Diplomatic Sentiment	$WDSI^{DE}$	Index	7,396	-0.943	0.829	-1.000	-3.000	2.000
India War-Related Diplomatic Sentiment	$WDSI^{IN}$	Index	9,600	-0.486	0.740	-0.429	-3.000	2.000
United Kingdom War-Related Diplomatic Sentiment	$WDSI^{UK}$	Index	5,816	-0.921	0.784	-0.857	-3.000	1.857
France War-Related Diplomatic Sentiment	$WDSI^{FR}$	Index	5,072	-0.493	0.981	0.000	-3.000	2.000
Italy War-Related Diplomatic Sentiment	$WDSI^{IT}$	Index	5,962	-0.440	0.639	-0.286	-2.857	2.571
Brazil War-Related Diplomatic Sentiment	$WDSI^{BR}$	Index	5,582	-0.483	0.795	0.000	-3.000	2.000
Canada War-Related Diplomatic Sentiment	$WDSI^{CA}$	Index	5,215	-0.579	0.868	-0.429	-3.000	2.000
Russia War-Related Diplomatic Sentiment	$WDSI^{RU}$	Index	5,877	-1.670	1.314	-2.000	-3.000	3.000
South Korea War-Related Diplomatic Sentiment	$WDSI^{KR}$	Index	9,233	-0.402	0.637	-0.286	-2.857	2.000
Australia War-Related Diplomatic Sentiment	$WDSI^{AU}$	Index	4,589	-0.309	0.904	0.000	-3.000	2.000
Spain War-Related Diplomatic Sentiment	$WDSI^{ES}$	Index	5,570	-0.476	1.045	0.000	-3.000	2.000
Mexico War-Related Diplomatic Sentiment	$WDSI^{MX}$	Index	3,814	-0.245	0.534	0.000	-3.000	1.429
Chinese Equity Market Return	R^{CN}	%	3,777	0.040	0.922	0.045	-1.812	1.918
U.S. Equity Market Return	R^{US}	%	3,777	0.061	0.837	0.080	-1.680	1.620
Developed Market Equity Return	R^{DEV}	%	3,777	0.033	0.698	0.060	-1.350	1.333
China Geopolitical Risk Index	GPR^{CN}	Index	3,777	0.635	0.255	0.577	0.281	1.247
U.S. Geopolitical Risk Index	GPR^{US}	Index	3,777	2.832	0.620	2.704	1.944	4.291
China Economic Policy Uncertainty	EPU^{CN}	Index	3,777	145.252	26.942	141.551	105.366	205.685
U.S. Economic Policy Uncertainty	EPU^{US}	Index	3,777	173.340	75.026	153.842	89.584	372.063
Gold Return	$Gold$	Log return	3,777	0.000	0.008	0.000	-0.016	0.016
RMB-USD Exchange Rate	FX	RMB/USD	3,777	6.666	0.354	6.655	6.133	7.249

Table 4: Baseline Results: Chinese War-Related Diplomatic Sentiment and Daily Stock Connect Flows

The table reports OLS estimates of Equation (1) with heteroskedasticity-robust standard errors in parentheses. The dependent variables are daily net northbound inflows (column 1), net southbound flows (column 2), and their combined net flow (column 3) through the Stock Connect programs. Flow variables are measured in hundred million USD. Equity returns are expressed in percentage points. All flow variables, equity returns, and control variables are winsorised at the 5th and 95th percentiles; untabulated winsorisation-threshold robustness checks using the 1st/99th and 10th/90th percentiles yield very similar coefficient estimates and significance levels. The three flow equations are estimated separately on dependent-variable-specific nonmissing samples, and the three flow measures are winsorised separately; as a result, the net-flow coefficient need not equal the inflow coefficient minus the outflow coefficient mechanically. Month fixed effects are included in all specifications; their coefficients are omitted for brevity. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) Inflow	(2) Outflow	(3) Net flow
$WDSI^{CN}$	0.009	-0.299***	0.424***
	(0.127)	(0.066)	(0.157)
R^{CN}	1.989***	0.394***	1.592***
	(0.113)	(0.061)	(0.142)
R^{US}	-0.213	0.151*	-0.396**
	(0.141)	(0.086)	(0.186)
$R_{Developed}$	0.912***	-0.335***	1.291***
	(0.185)	(0.110)	(0.240)
GPR^{CN}	2.318***	-0.786***	3.102***
	(0.465)	(0.286)	(0.622)
GPR^{US}	-1.141***	0.148	-1.195***
	(0.200)	(0.119)	(0.266)
EPU^{CN}	-0.001	-0.002	-0.001
	(0.004)	(0.002)	(0.006)
EPU^{US}	-0.000	0.008***	-0.009***
	(0.002)	(0.001)	(0.003)
$Gold$	-5.142	5.310	-16.010
	(13.107)	(7.701)	(17.415)
FX	-1.133***	2.011***	-2.652***
	(0.386)	(0.228)	(0.514)
Month-of-year FE	Yes	Yes	Yes
Observations	2,205	2,224	2,190
R^2	0.222	0.129	0.137

Table 5: Robustness to Alternative Smoothing Windows for Chinese War-Related Sentiment

The table reports OLS estimates of Equation (1) with heteroskedasticity-robust standard errors in parentheses. The dependent variables are daily net northbound inflows (Inflow), net southbound flows (Outflow), and their combined net flow (Net flow) through the Stock Connect programs, measured in hundred million USD. Columns (1)–(3) use a three-day moving average of the Chinese war-related diplomatic sentiment index $WDSI_t^{CN,3}$, while columns (4)–(6) use a thirty-day moving average $WDSI_t^{CN,30}$. Equity returns are expressed in percentage points. All flow variables, returns, and controls are winsorised at the 5th and 95th percentiles. Month fixed effects are included in all specifications but their coefficients are omitted for brevity. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	3-day average			30-day average		
	(1) Inflow	(2) Outflow	(3) Net flow	(4) Inflow	(5) Outflow	(6) Net flow
$WDSI^{CN}$	0.016	-0.275***	0.364***	0.168	-0.669***	0.977***
	(0.109)	(0.061)	(0.140)	(0.199)	(0.096)	(0.240)
R^{CN}	1.989***	0.395***	1.590***	1.991***	0.388***	1.603***
	(0.113)	(0.061)	(0.142)	(0.113)	(0.060)	(0.142)
R^{US}	-0.213	0.145*	-0.388**	-0.212	0.150*	-0.395**
	(0.141)	(0.087)	(0.186)	(0.141)	(0.085)	(0.185)
$R_{Developed}$	0.912***	-0.325***	1.277***	0.914***	-0.339***	1.298***
	(0.185)	(0.110)	(0.241)	(0.184)	(0.109)	(0.239)
GPR^{CN}	2.317***	-0.759***	3.064***	2.344***	-0.866***	3.231***
	(0.466)	(0.286)	(0.623)	(0.463)	(0.288)	(0.623)
GPR^{US}	-1.140***	0.144	-1.194***	-1.123***	0.114	-1.145***
	(0.200)	(0.119)	(0.266)	(0.201)	(0.119)	(0.267)
EPU^{CN}	-0.001	-0.002	-0.001	-0.001	-0.000	-0.003
	(0.004)	(0.002)	(0.006)	(0.004)	(0.002)	(0.006)
EPU^{US}	-0.000	0.008***	-0.009***	-0.000	0.008***	-0.009***
	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)
$Gold$	-5.119	5.003	-15.575	-5.060	5.110	-15.674
	(13.096)	(7.700)	(17.400)	(13.096)	(7.657)	(17.380)
FX	-1.129***	2.007***	-2.660***	-1.073***	1.883***	-2.460***
	(0.385)	(0.229)	(0.514)	(0.390)	(0.229)	(0.517)
Month-of-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,205	2,224	2,190	2,205	2,224	2,190
R^2	0.222	0.130	0.137	0.222	0.138	0.141

Table 6: VAR Dynamic Robustness: Daily and Weekly VAR(1) Specifications

The table reports residualised VAR(1) results for the three Stock Connect flow channels. Panel A uses daily $WDSI_t^{CN}$ (seven-day moving average), Panel B uses daily $WDSI_t^{CN,30}$ (thirty-day moving average), Panel C uses weekly $WDSI_t^{CN,30}$, and Panel D uses weekly $WDSI_t^{CN}$. Weekly observations are built from Monday-anchored weeks, with flow variables summed and sentiment/control variables averaged within week. For each channel, both the sentiment index and the flow variable are first residualised on the full baseline control set and month-of-year fixed effects from Equation (1), and the VAR is then estimated on the residualised series. Coefficients are reported with standard errors in parentheses. Granger tests are one-lag tests. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Inflow / Sentiment	Outflow / Sentiment	Net flow / Sentiment
Panel A. Daily VAR(1), baseline $WDSI_t^{CN}$ (7-day MA)			
$L1.$ Sentiment	-0.002	-0.041**	0.031
	(0.021)	(0.020)	(0.020)
$L1.$ Flow	0.213***	0.372***	0.283***
	(0.021)	(0.020)	(0.021)
p(Sentiment → Flow)	0.932	0.036	0.125
p(Flow → Sentiment)	0.512	0.346	0.990
Observations	2,204	2,223	2,189
Panel B. Daily VAR(1), alternative $WDSI_t^{CN,30}$ (30-day MA)			
$L1.$ Sentiment	0.018	-0.072***	0.061***
	(0.021)	(0.020)	(0.021)
$L1.$ Flow	0.213***	0.366***	0.280***
	(0.021)	(0.020)	(0.021)
p(Sentiment → Flow)	0.385	0.000	0.003
p(Flow → Sentiment)	0.979	0.133	0.384
Observations	2,204	2,223	2,189
Panel C. Weekly VAR(1), alternative $WDSI_t^{CN,30}$ (30-day MA)			
$L1.$ Sentiment	0.048	-0.079**	0.105**
	(0.044)	(0.038)	(0.041)
$L1.$ Flow	0.207***	0.533***	0.398***
	(0.044)	(0.038)	(0.041)
p(Sentiment → Flow)	0.278	0.040	0.011
p(Flow → Sentiment)	0.451	0.010	0.415
Observations	497	496	496
Panel D. Weekly VAR(1), baseline $WDSI_t^{CN}$ (7-day MA)			
$L1.$ Sentiment	0.030	-0.091**	0.080*
	(0.044)	(0.038)	(0.041)
$L1.$ Flow	0.208***	0.537***	0.406***
	(0.044)	(0.038)	(0.041)
p(Sentiment → Flow)	0.498	0.016	0.052
p(Flow → Sentiment)	0.859	0.012	0.341
Observations	497	496	496

Table 7: Placebo Sentiment Measures: Chinese Stock Connect Flows

This table reports OLS estimates of Equation (1) with heteroskedasticity-robust standard errors in parentheses. The dependent variables are daily net northbound inflows (Inflow), net southbound flows (Outflow), and their combined net flow (Net flow) through the Stock Connect programs, measured in hundred million USD. Columns (1)–(3) replace the baseline war-related sentiment index $WDSI_t^{CN}$ with an economic-related sentiment index $EDSI_t^{CN}$; columns (4)–(6) use an index of other diplomatic sentiment $ODSI_t^{CN}$; columns (7)–(9) use the negative-part war sentiment index $WDSI_t^{CN,-}$, which equals $WDSI_t^{CN}$ when the index is negative and zero otherwise. All sentiment indices are expressed by their respective seven-day moving averages. Equity returns are expressed in percentage points. All flow variables, returns, and controls are winsorised at the 5th and 95th percentiles. Month fixed effects are included in all specifications but their coefficients are omitted for brevity. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Economic conflict sentiment			Other diplomatic sentiment			War sentiment, negative only		
	(1) Inflow	(2) Outflow	(3) Net flow	(4) Inflow	(5) Outflow	(6) Net flow	(7) Inflow	(8) Outflow	(9) Net flow
Sentiment index	0.028	-0.012	-0.046	-0.128	-0.002	-0.098	-0.041	-0.406***	0.515***
	(0.138)	(0.080)	(0.181)	(0.116)	(0.065)	(0.152)	(0.146)	(0.083)	(0.191)
R^{CN}	1.988***	0.398***	1.588***	1.986***	0.397***	1.585***	1.988***	0.397***	1.588***
	(0.113)	(0.061)	(0.143)	(0.113)	(0.061)	(0.143)	(0.113)	(0.060)	(0.143)
R^{US}	-0.212	0.152*	-0.402**	-0.214	0.153*	-0.401**	-0.213	0.153*	-0.399**
	(0.141)	(0.087)	(0.187)	(0.142)	(0.087)	(0.187)	(0.142)	(0.086)	(0.186)
$R_{Developed}$	0.912***	-0.332***	1.285***	0.918***	-0.332***	1.290***	0.911***	-0.335***	1.291***
	(0.185)	(0.110)	(0.242)	(0.185)	(0.110)	(0.242)	(0.185)	(0.110)	(0.241)
GPR^{CN}	2.327***	-0.778***	3.066***	2.312***	-0.775***	3.074***	2.316***	-0.792***	3.106***
	(0.470)	(0.291)	(0.629)	(0.466)	(0.287)	(0.623)	(0.466)	(0.286)	(0.622)
GPR^{US}	-1.142***	0.197*	-1.265***	-1.116***	0.198*	-1.244***	-1.149***	0.139	-1.189***
	(0.198)	(0.118)	(0.263)	(0.199)	(0.117)	(0.263)	(0.200)	(0.119)	(0.266)
EPU^{CN}	-0.001	-0.003	0.001	-0.000	-0.003	0.002	-0.001	-0.003	0.000
	(0.004)	(0.002)	(0.005)	(0.004)	(0.002)	(0.005)	(0.004)	(0.002)	(0.005)
EPU^{US}	-0.000	0.009***	-0.010***	-0.001	0.009***	-0.010***	-0.000	0.008***	-0.009***
	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)
$Gold$	-5.173	5.315	-16.012	-5.388	5.299	-16.244	-5.151	5.254	-15.952
	(13.116)	(7.740)	(17.471)	(13.089)	(7.738)	(17.444)	(13.107)	(7.688)	(17.418)
FX	-1.127***	2.139***	-2.851***	-1.165***	2.143***	-2.856***	-1.148***	2.026***	-2.691***
	(0.379)	(0.225)	(0.504)	(0.381)	(0.227)	(0.507)	(0.383)	(0.226)	(0.509)
Month-of-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,205	2,224	2,190	2,205	2,224	2,190	2,205	2,224	2,190
R^2	0.222	0.123	0.135	0.222	0.123	0.135	0.222	0.131	0.138

Table 8: Endogeneity Checks: Reverse Causality and Detrended Sentiment

Panel A reports reverse-causality regressions in which the dependent variable is the one-day-ahead Chinese war-related diplomatic sentiment index $WDSI_{t+1}^{CN}$ and the key regressors are contemporaneous and strictly calendar-day-lagged Stock Connect flows (inflow, outflow, or net flow). The lagged flow is set to missing when the previous calendar day falls outside the Stock Connect trading calendar. Panel B re-estimates Equation (1) using a detrended sentiment index $WDSI_t^{CN}$ – the residuals from regressing $WDSI_t^{CN}$ on a linear time trend – in place of the raw index. All specifications include the same set of controls as in Table 4 as well as month-of-year fixed effects; their coefficients are omitted for brevity. Robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Panel A. Future sentiment			Panel B. Detrended sentiment		
	(1) Inflow	(2) Outflow	(3) Net flow	(4) Inflow	(5) Outflow	(6) Net flow
$Flow_{k,t}^{CN}$	-0.00811	0.00417	-0.00329			
	(0.00664)	(0.01096)	(0.00507)			
$Flow_{k,t-1}^{CN}$	0.00484	-0.01542	0.00531			
	(0.00579)	(0.01073)	(0.00461)			
$WDSI_t^{CN}$				0.025	-0.231***	0.377**
(detrended)				(0.127)	(0.066)	(0.157)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,734	1,749	1,719	2,205	2,224	2,190
R^2	0.051	0.051	0.051	0.222	0.127	0.137

Table 9: Dynamic Specifications with Lagged Flows and Controls

The table augments the baseline Stock Connect regressions in Equation (1) with lagged dependent variables. Panel A includes the first lag of each flow variable; Panel B includes both first and second lags. All specifications retain the full set of controls as in Table 4 as well as month-of-year fixed effects; coefficients on these controls and on lagged controls are omitted for brevity. Robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Panel A. First-order lags			Panel B. Second-order lags		
	(1) Inflow	(2) Outflow	(3) Net flow	(4) Inflow	(5) Outflow	(6) Net flow
$WDSI_t^{CN}$	0.032	-0.211***	0.360**	0.082	-0.159***	0.366**
	(0.121)	(0.058)	(0.146)	(0.121)	(0.056)	(0.143)
$Flow_{k,t-1}^{CN}$	0.210***	0.378***	0.285***	0.178***	0.279***	0.230***
	(0.025)	(0.027)	(0.025)	(0.026)	(0.029)	(0.026)
$Flow_{k,t-2}^{CN}$				0.132***	0.272***	0.179***
				(0.024)	(0.028)	(0.026)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls($t-1$)	Yes	Yes	Yes	Yes	Yes	Yes
Controls($t-2$)	No	No	No	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,204	2,223	2,189	2,203	2,222	2,188
R^2	0.287	0.277	0.251	0.328	0.334	0.295

Table 10: Instrumental Variable Estimates

Column (1) reports the first-stage regression of the Chinese war-related diplomatic sentiment index $WDSI_t^{CN}$ on the spokesperson fixed-sentiment measure, defined as the average sentiment across all answers provided by each spokesperson, along with the full set of controls used in Table 4 and month-of-year fixed effects. Columns (2)–(4) report second-stage 2SLS estimates of Equation (1), where $WDSI_t^{CN}$ is instrumented with the spokesperson fixed-sentiment variable. The dependent variables are daily net northbound inflows (column 2), net southbound flows (column 3), and their combined net flow (column 4) through the Stock Connect programs. All specifications include the same controls and fixed effects as the baseline OLS regressions; their coefficients are omitted for brevity. Robust standard errors are reported in parentheses. "First-stage Wald χ^2 (excluded IV)" reports the robust first-stage relevance test in each equation sample. "KP rk Wald F" and "KP rk LM p-value" are the Kleibergen–Paap weak-identification and under-identification diagnostics (from `ivreg2`, robust option). *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) First stage	(2) Inflow	(3) Outflow	(4) Net flow
Dependent variable	$WDSI_t^{CN}$	$Flow_{\text{Inflow},t}^{CN}$	$Flow_{\text{Outflow},t}^{CN}$	$Flow_{\text{Net},t}^{CN}$
Spokesperson fixed sentiment	0.919***			
	(0.067)			
$WDSI_t^{CN}$		0.418	-2.444***	2.770***
		(0.512)	(0.347)	(0.708)
Controls	Yes	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes	Yes
First-stage Wald χ^2 (excluded IV)	186.314	186.314	186.341	182.255
KP rk Wald F		188.192	188.203	184.105
KP rk LM p-value		< 0.001	< 0.001	< 0.001
Observations	2,205	2,205	2,224	2,190

Table 11: War-Related Sentiment in 14 Foreign Countries and Chinese Stock Connect Flows

The table reports full-sample daily OLS estimates with heteroskedasticity-robust standard errors in parentheses. Columns correspond to Chinese Stock Connect flow equations (inflow, outflow, net flow), and rows report coefficients on each foreign war-related sentiment index. Dependent variables are Chinese Stock Connect flows measured in hundred million USD. The sample runs from November 17, 2014 to August 16, 2024 for all countries except Mexico, whose sample begins on November 9, 2015 because of later index availability. Most country equations therefore use 2,205, 2,224, and 2,190 observations in the inflow, outflow, and net-flow specifications, respectively, while Mexico uses 1,986, 2,000, and 1,972 observations. All specifications include the same control block (Chinese/U.S./developed-market returns, China/U.S. GPR and EPU, gold returns, RMB-USD exchange rates) and month-of-year fixed effects. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) Inflow	(2) Outflow	(3) Net flow
$WDSI^{US}$	0.040	-0.252**	0.309
	(0.161)	(0.105)	(0.219)
$WDSI^{UK}$	0.108	-0.246***	0.328**
	(0.120)	(0.074)	(0.163)
$WDSI^{JP}$	0.046	-0.184**	0.294*
	(0.124)	(0.078)	(0.166)
$WDSI^{KR}$	0.302**	-0.019	0.295*
	(0.129)	(0.080)	(0.169)
$WDSI^{AU}$	-0.203*	0.053	-0.228
	(0.106)	(0.066)	(0.140)
$WDSI^{BR}$	0.219**	-0.113*	0.376***
	(0.108)	(0.064)	(0.143)
$WDSI^{CA}$	0.043	-0.100	0.158
	(0.117)	(0.073)	(0.157)
$WDSI^{DE}$	0.087	-0.371***	0.490***
	(0.134)	(0.083)	(0.181)
$WDSI^{ES}$	0.171	0.069	0.074
	(0.108)	(0.071)	(0.148)
$WDSI^{FR}$	0.001	-0.195**	0.237
	(0.142)	(0.084)	(0.190)
$WDSI^{IN}$	-0.019	0.288***	-0.341*
	(0.142)	(0.092)	(0.204)
$WDSI^{IT}$	-0.262	-0.192**	-0.141
	(0.165)	(0.098)	(0.224)
$WDSI^{MX}$	0.210	-0.309**	0.583**
	(0.191)	(0.125)	(0.252)
$WDSI^{RU}$	0.099	-0.220***	0.371***
	(0.077)	(0.045)	(0.100)
Controls	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes
Observations	2,205	2,224	2,190

Table 12: Domestic War-Related Sentiment and Own-Country International Equity Positions: Monthly Mean Specifications

The table reports monthly OLS estimates with heteroskedasticity-robust standard errors in parentheses. Panel A uses the monthly mean of the daily seven-day moving-average WDSI series, and Panel B uses the monthly mean of the daily thirty-day moving-average WDSI series. These are full-signed WDSI measures that retain both positive and negative daily values; they are not raw daily scores, monthly low-points, or negative-part transformations. The core rows (United States, Japan, South Korea, Brazil, and Canada) use broad monthly own-country equity-flow equations. The France row uses a monthly-updated 12-month cumulative portfolio-investment balance proxy, the Italy row uses monthly portfolio-investment assets/liabilities totals from Banca d'Italia Table 5, and the Russia row uses the monthly stock of OFZ held by non-residents as an official proxy for foreign portfolio demand. Blanks indicate that the monthly proxy does not provide a distinct inflow/outflow decomposition. All specifications include country-level controls (U.S./developed-market returns, local GPR, local EPU, local exchange-rate control) and month-of-year fixed effects. Country-specific monthly sample sizes are reported in parentheses in the row labels. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Panel A. Monthly mean of 7-day WDSI			Panel B. Monthly mean of 30-day WDSI		
	(1) Inflow$_{own}$	(2) Outflow$_{own}$	(3) Net flow$_{own}$	(4) Inflow$_{own}$	(5) Outflow$_{own}$	(6) Net flow$_{own}$
$WDSI^{US}$	-0.232***	-0.126**	-15.118***	-0.282***	-0.152***	-18.214***
(N=132)	(0.082)	(0.051)	(2.510)	(0.086)	(0.055)	(2.568)
$WDSI^{JP}$	0.038	0.032	4.836	0.025	0.019	4.234
(N=132)	(0.043)	(0.043)	(3.800)	(0.060)	(0.060)	(4.255)
$WDSI^{KR}$	0.137*	0.129*	8.575***	0.169**	0.161**	8.532***
(N=131)	(0.076)	(0.075)	(2.358)	(0.081)	(0.080)	(2.691)
$WDSI^{BR}$	0.033	0.032	-0.324	0.053	0.055	-0.125
(N=179)	(0.037)	(0.039)	(1.106)	(0.040)	(0.040)	(1.200)
$WDSI^{CA}$	0.552	0.605	0.350	1.249	0.324	0.652
(N=167)	(1.421)	(1.512)	(1.611)	(1.532)	(1.654)	(1.676)
$WDSI^{FR}$ proxy	–	–	1.965	–	–	2.315
(N=50)			(1.630)			(1.683)
$WDSI^{IT}$ proxy	11.388	-25.631***	2.941	5.654	-15.197	-7.785
(N=22)	(9.159)	(5.985)	(5.949)	(5.812)	(10.743)	(8.695)
$WDSI^{RU}$ proxy	–	–	0.126***	–	–	0.147***
(N=167)			(0.031)			(0.031)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 13: Chinese War-Related Sentiment and the A-H Share Premium

The table reports OLS estimates using the same control block and month-of-year fixed effects as Equation (1), with the A-H share premium as the dependent variable. The key explanatory variable is the Chinese war-related sentiment index $WDSI^{CN}$ (seven-day moving average). Column (1) includes only contemporaneous regressors; columns (2) and (3) additionally include one and two lags of the A-H premium, respectively, to capture persistence. Columns (1)–(3) use the level of the A-H premium as the dependent variable, whereas column (4) uses its first difference and includes one and two lags of the differenced dependent variable. The A-H premium is measured in percentage points. Standard errors are heteroskedasticity-robust and reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) A-H Share Premium	(2) A-H Share Premium	(3) A-H Share Premium	(4) Δ A-H Share Premium
$WDSI^{CN}$	-0.887***	-0.071*	-0.075*	-0.075*
	(0.296)	(0.041)	(0.041)	(0.041)
A-H Share Premium$_{t-1}$		0.994***	0.908***	
		(0.002)	(0.036)	
A-H Share Premium$_{t-2}$			0.086**	
			(0.036)	
Δ A-H Share Premium$_{t-1}$				-0.094***
				(0.036)
Δ A-H Share Premium$_{t-2}$				-0.065**
				(0.027)
R^{CN}	-0.119	0.018	0.012	0.010
	(0.239)	(0.034)	(0.034)	(0.034)
R^{US}	0.136	0.154***	0.153***	0.153***
	(0.295)	(0.041)	(0.041)	(0.041)
$R_{Developed}$	-0.460	-0.510***	-0.510***	-0.511***
	(0.390)	(0.054)	(0.054)	(0.055)
GPR^{CN}	6.539***	0.047	0.047	0.009
	(0.893)	(0.113)	(0.113)	(0.110)
GPR^{US}	3.150***	0.053	0.052	0.034
	(0.415)	(0.042)	(0.042)	(0.043)
EPU^{CN}	0.027***	-0.000	-0.000	-0.000
	(0.008)	(0.001)	(0.001)	(0.001)
EPU^{US}	-0.017***	-0.000	-0.000	-0.000
	(0.003)	(0.000)	(0.000)	(0.000)
$Gold$	20.322	9.634***	9.735***	9.675***
	(25.941)	(3.177)	(3.175)	(3.177)
FX	21.097***	0.059	0.040	-0.089
	(0.681)	(0.087)	(0.088)	(0.073)
Month-of-year FE	Yes	Yes	Yes	Yes
Observations	3,401	3,400	3,399	3,398
R^2	0.390	0.993	0.993	0.060

Table 14: Chinese War-Related Sentiment and Equity-Index Returns

The table reports OLS regressions of Chinese equity-index returns (in percentage points) on war-related diplomatic sentiment. Panel A uses daily baseline sentiment $WDSI^{CN}_{t-1}$ (seven-day moving average, one-day lag). Panel B uses weekly baseline sentiment $WDSI^{CN}_{t-1}$ (seven-day moving average, lag 1). Panel C uses weekly alternative sentiment $WDSI^{CN,30}_{t-1}$ (thirty-day moving average, lag 1). Panel D uses monthly contemporaneous sentiment $WDSI^{CN,30}_{t}$. All specifications include the same control block as Equation (1): U.S./developed-market returns, China/U.S. GPR, China/U.S. EPU, gold, FX, and month-of-year fixed effects. Robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) A-share broad	(2) CSI 300	(3) CSI 500	(4) CSI 1000
Panel A. Daily baseline $WDSI^{CN}_{t-1}$ (7-day MA, lag 1)				
$WDSI^{CN}_{t-1}$	-0.040	-0.041	-0.040	-0.067
	(0.033)	(0.035)	(0.052)	(0.044)
Controls	Yes	Yes	Yes	Yes
Observations	3,777	3,777	2,924	3,777
R^2	0.087	0.091	0.065	0.056
Panel B. Weekly baseline $WDSI^{CN}_{t-1}$ (7-day MA)				
$WDSI^{CN}_{t-1}$	-0.297*	-0.263*	-0.300	-0.430*
	(0.155)	(0.159)	(0.253)	(0.231)
Controls	Yes	Yes	Yes	Yes
Observations	824	825	639	824
R^2	0.130	0.134	0.103	0.093
Panel C. Weekly alternative $WDSI^{CN,30}_{t-1}$ (30-day MA)				
$WDSI^{CN,30}_{t-1}$	-0.461**	-0.393*	-0.707*	-0.634*
	(0.205)	(0.215)	(0.377)	(0.324)
Controls	Yes	Yes	Yes	Yes
Observations	822	823	639	822
R^2	0.130	0.132	0.111	0.093
Panel D. Monthly $WDSI^{CN,30}_{t}$ (contemporaneous)				
$WDSI^{CN,30}_{t}$	-1.917***	-1.769**	-2.473***	-2.774***
	(0.571)	(0.695)	(0.930)	(0.960)
Controls	Yes	Yes	Yes	Yes
Observations	194	194	150	194
R^2	0.285	0.278	0.325	0.274

Table 15: War-Related Sentiment and Market Activity

The table reports OLS estimates of Equation (1) with measures of market activity as dependent variables. Panel A uses the log of daily trading volume for the CSI 300, CSI 1000, Nasdaq, and Dow Jones Industrial Average. Panel B uses daily realized volatility of the same indices as dependent variables. In columns (1), (2), (5), and (6) the key explanatory variable is the Chinese war-related sentiment index $WDSI^{CN}$ (seven-day moving average); in columns (3), (4), (7), and (8) it is the U.S. war-related sentiment index $WDSI^{US}$. Standard errors are heteroskedasticity-robust and reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Log trading volume				
	(1) CSI 300	(2) CSI 1000	(3) Nasdaq	(4) Dow Jones Industrial Average
$WDSI^{CN}$	-0.053***	-0.107***		
	(0.014)	(0.016)		
$WDSI^{US}$			-0.144***	-0.112***
			(0.009)	(0.012)
Controls	Yes	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes	Yes
Observations	3,777	3,777	3,690	3,711
R^2	0.173	0.358	0.491	0.617
Panel B. Realized volatility				
	(5) CSI 300	(6) CSI 1000	(7) Nasdaq	(8) Dow Jones Industrial Average
$WDSI^{CN}$	0.000014***	0.000030***		
	(0.000005)	(0.000008)		
$WDSI^{US}$			-0.000028***	-0.000011***
			(0.000005)	(0.000003)
Controls	Yes	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes	Yes
Observations	3,897	3,897	3,690	3,711
R^2	0.044	0.035	0.093	0.086

Online Supplement

When Words Move Money: Diplomatic Sentiment and International Capital Flows

Supplementary Materials

I. Variable Definitions

Table I.1 (and its continuation) summarises the construction and definitions of variables used in the empirical analysis. Panel A reports diplomatic-sentiment measures. Panel B reports Stock Connect and international-flow variables. Panel C collects price, return, activity, and volatility variables. Panel D defines macro-financial controls. Panel E lists econometric transformations and reported diagnostics. Unless otherwise noted, flow variables, returns, and controls are winsorised at the 5th and 95th percentiles, and higher WDSI values indicate more positive war-related diplomatic sentiment.

Table I.1: Variable Definitions

Variable	Definition
Panel A: Diplomatic-sentiment variables	
$WDSI_t^c$	Seven-day moving average of country c's daily war-related diplomatic sentiment score, scaled from -3 to 3.
$WDSI_t^{c,3}$	Three-day moving average of country c's WDSI.
$WDSI_t^{c,30}$	Thirty-day moving average of country c's WDSI.
Country code c	CN, US, JP, DE, IN, UK, FR, IT, BR, CA, RU, KR, AU, ES, and MX denote the 15 sample economies.
$EDSI_t^c$	Economic- and trade-related diplomatic sentiment index for country c.
$ODSI_t^c$	Diplomatic sentiment index for non-war and non-economic topics.
$WDSI_t^{CN,-}$	Negative part of China's WDSI, equal to $WDSI_t^{CN}$ when negative and zero otherwise.
Detrended $WDSI_t^{CN}$	Residual from regressing China's baseline WDSI on a linear time trend.
$WDSI_{t+1}^{CN}$	One-calendar-day lead of China's WDSI.
$WDSI_m^c$, $WDSI_m^{c,30}$	Monthly aggregation of country c's daily WDSI series.
$WDSI_m^{CN,\min}$	Monthly low-point Chinese WDSI used in the GPR benchmark VAR.
Spokesperson fixed sentiment	Seven-day moving average of spokesperson-level sentiment.
Panel B: Stock Connect and international-flow variables	
$Flow_{k,t}^{CN}$	Chinese daily Stock Connect flow for channel k, measured in hundred million U.S. dollars.
Inflow	Net northbound purchases of mainland A-shares.
Outflow	Net southbound purchases of Hong Kong-listed equities by mainland investors.
NetFlow	Net northbound inflow minus net southbound outflow.
$Flow_{k,t-1}^{CN}$, $Flow_{k,t-2}^{CN}$	One- and two-period lags of Stock Connect flow channel k.
Inflow_{own}	Country-level international equity inflow.
Outflow_{own}	Country-level international equity outflow.
Net flow_{own}	Country-level net international equity flow or position measure.
Net flow_{own}^{FR}	France's monthly 12-month cumulative portfolio-investment balance.
Inflow_{own}^{IT}, $\text{Outflow}_{own}^{IT}$	Italy's monthly portfolio-investment assets and liabilities totals.
OFZ non-resident holdings	Monthly stock of Russian OFZ government bonds held by non-residents.
Target-country flow	Monthly target-country flow outcome in the pairwise network regressions.
Panel C: Price, return, activity, and volatility variables	
R^{CN}	Chinese A-share market return, in percentage points.
R^{US}	U.S. equity market return, in percentage points.
R^{DEV}, $R_{Developed}$	Developed-market equity return excluding the United States.
Chinese equity-index returns	Returns on the A-share broad market, CSI 300, CSI 500, and CSI 1000 indices.
Nasdaq; DJIA	U.S. equity indices used in activity and volatility tests.
A-H Share Premium	Percentage price premium of mainland A-shares over matched Hong Kong-listed H-shares.
Δ A-H Share Premium	First difference of the A-H share premium.
Lagged A-H premium	One- and two-day lags of the A-H share premium.
Lagged Δ A-H premium	One- and two-day lags of Δ A-H Share Premium.
log(Volume)	Natural logarithm of daily trading volume.
Realized volatility	Daily realized volatility of the corresponding equity index.

Table I.1: Variable Definitions (Continued)

Variable	Definition
Panel D: Macro-financial control variables	
$\mathbf{R}_t$	Vector of equity-return controls.
$\mathbf{Z}_t$	Vector of non-return macro-financial controls.
GPR^{CN}	China-specific geopolitical risk index.
GPR^{US}	U.S.-specific geopolitical risk index.
Local GPR	Country-specific geopolitical risk control.
EPU^{CN}	China economic policy uncertainty index.
EPU^{US}	U.S. economic policy uncertainty index.
Local EPU	Country-specific economic policy uncertainty control.
Gold	Log first difference of gold prices.
FX	Bilateral RMB per U.S. dollar exchange rate.
Local FX	Country-level exchange-rate control.
$\mu_{m(t)}$	Month-of-year fixed effects (January–December indicators); excludes calendar year-month fixed effects.
$\varepsilon^{CN}_{k,t}$, $u_{k,t}$	Regression error terms.
Panel E: Econometric transformations and reported diagnostics	
L1. Sentiment; L1. Flow	First lag of residualised sentiment and flow variables in VARs.
p(Sentiment $\rightarrow$ Flow)	Granger-test p-value for sentiment predicting flow.
p(Flow $\rightarrow$ Sentiment)	Granger-test p-value for flow predicting sentiment.
p-value: WDSI$\rightarrow$GPR	Granger-test p-value for WDSI predicting GPR.
p-value: GPR$\rightarrow$WDSI	Granger-test p-value for GPR predicting WDSI.
p-value: A-H$\rightarrow$Flow	Granger-test p-value for A-H premium predicting Stock Connect flow.
p-value: Flow$\rightarrow$A-H	Granger-test p-value for Stock Connect flow predicting A-H premium.
Lag order p	Number of lags in a VAR specification.
AIC	Akaike information criterion.
BIC	Bayesian information criterion.
First-stage Wald χ^2	Robust relevance test for the excluded instrument.
KP rk Wald F	Kleibergen–Paap weak-identification statistic.
KP rk LM p-value	Kleibergen–Paap under-identification test p-value.
R^2; Observations	Regression fit and usable sample size.

II. Data Sources

Table II.1 reports the official diplomatic text corpora used to construct the WDSI database for the 15 largest world economies. The document counts and coverage windows refer to the validated WDSI input as of 20 April 2026, with coverage measured by publication date. Historical material outside the reported modern windows is excluded from Table 3 and from the cross-country empirical results. The table therefore describes the raw measurement input, while the estimation sample in the main text imposes additional restrictions from financial-market data availability and regression design.

A. Chinese Ministry of Foreign Affairs Press Conferences

We construct the Chinese WDSI from transcripts of regular press conferences held by China's Ministry of Foreign Affairs between 5 January 2010 and 17 April 2026. Each press conference is an official Q&A briefing in which a spokesperson responds to questions from domestic and foreign journalists. We merge all Q&A units released on the same calendar day into one daily document, so the empirical unit is a daily MFA text at the publication-day level. The resulting series contains several thousand daily observations and provides near-weekday coverage, with fewer observations around major holidays. Because the transcripts are published on the MFA's official website and disseminated through state media, they are authoritative statements of China's diplomatic position on current international affairs.

The Chinese press-conference corpus has three features that make it useful for measuring war-related diplomatic sentiment. First, because each session contains a limited number of questions, journalists tend to ask about the most salient war, conflict, and security issues of the day. The text therefore loads heavily on contemporaneous geopolitical shocks, including regional wars, military confrontations, security threats, and other high-profile incidents. Second, spokesperson responses are scripted and cleared within the foreign-policy bureaucracy, so their tone reflects formal state positions and avoids off-the-cuff commentary. For example, at the 9 December 2022 press conference, MFA Spokesperson Mao Ning responded to a question on U.S. military aid to Taiwan by "firmly opposing" the legislation, urging the United States to "stop playing the Taiwan card", and warning that China would "take all necessary measures to firmly safeguard its sovereignty". As shown in Table 1, Panel B, the pipeline classifies this exchange as high-intensity negative sentiment and assigns the lowest score of -3. Third,

Table II.1: Diplomatic Text Corpora Used to Construct the WDSI

The table summarises the official diplomatic text corpora currently used to construct the WDSI database for the 15 largest world economies. All texts are collected from official foreign-ministry or foreign-office websites and archival pages.

Countries	Text source	Issuing institution	Time coverage	Number of texts
China ($WDSI^{CN}$)	Regular press-conference transcripts	Ministry of Foreign Affairs of the PRC	2010-01-05 to 2026-04-17	3,095 daily texts
United States ($WDSI^{US}$)	Press releases, briefings, remarks, readouts and announcements	U.S. Department of State	2010-01-01 to 2026-04-18	35,329 documents
United Kingdom ($WDSI^{UK}$)	Press releases, speeches, oral statements and authored articles	UK Foreign, Commonwealth and Development Office	2010-05-12 to 2026-04-19	17,844 documents
Japan ($WDSI^{JP}$)	Press conferences and written statements	Ministry of Foreign Affairs of Japan	2013-03-01 to 2026-04-20	14,581 documents
South Korea ($WDSI^{KR}$)	Press briefings, press releases, ministerial remarks and related MFA texts	Ministry of Foreign Affairs of the Republic of Korea	2001-01-03 to 2026-04-20	13,262 documents
Australia ($WDSI^{AU}$)	Media releases and ministerial statements	Department of Foreign Affairs and Trade of Australia	2013-09-18 to 2026-04-17	2,870 documents
Brazil ($WDSI^{BR}$)	Press releases, statements and joint communiques	Ministry of Foreign Affairs of Brazil (Itamaraty)	2011-01-01 to 2026-04-19	3,642 documents
Canada ($WDSI^{CA}$)	News releases, statements, readouts, speeches and media advisories	Global Affairs Canada	2012-01-01 to 2026-04-17	9,201 documents
Germany ($WDSI^{DE}$)	Press releases, speeches, interviews and related articles	Federal Foreign Office of Germany	2006-01-01 to 2026-04-16	11,863 documents
Spain ($WDSI^{ES}$)	Communiques, press notes and official statements	Ministry of Foreign Affairs, European Union and Cooperation of Spain	2011-01-12 to 2026-04-18	2,492 documents
France ($WDSI^{FR}$)	Press statements, speeches, live Q&A and policy-event interventions	Ministry for Europe and Foreign Affairs of France	2012-05-24 to 2026-04-18	2,476 documents
India ($WDSI^{IN}$)	Press releases, media briefings, statements, interviews and parliamentary answers	Ministry of External Affairs of India	2000-01-01 to 2026-04-19	21,380 documents
Italy ($WDSI^{IT}$)	Press releases, statements, interviews and related articles	Ministry of Foreign Affairs and International Cooperation of Italy	2009-12-18 to 2026-04-20	10,002 documents
Mexico ($WDSI^{MX}$)	Press releases and joint statements	Secretariat of Foreign Affairs of Mexico	2015-11-03 to 2026-04-18	4,806 documents
Russia ($WDSI^{RU}$)	Briefings, news releases and official statements	Ministry of Foreign Affairs of the Russian Federation	2010-03-08 to 2026-04-15	1,342 documents

these briefings typically occur in the afternoon before Asian markets close and are triggered by news events, with no evident dependence on market outcomes. This timing strengthens the interpretation of the resulting WDSI as an official policy signal for the flow and market-response tests that follow.

B. United States Department of State Official Releases

For the United States, we construct a parallel corpus of official releases issued by the U.S. Department of State between 1 January 2010 and 18 April 2026, including press releases, briefings, remarks, readouts, announcements, and related official statements. These texts are formal announcements or comments, usually issued in the name of the Secretary of State or the Department Spokesperson, that state the U.S. government's position on international developments. Each document is a stand-alone release, typically prompted by discrete events such as conflict outbreaks, military operations, terrorist attacks, peace agreements, or other security-related developments. The daily frequency ranges from zero to several releases because the corpus is event-driven. The corpus covers a broad set of war and security topics consistent with the global reach of U.S. foreign policy and often expresses condemnation, concern, support for allies, or calls for de-escalation. We process all releases through the same multi-stage evaluation pipeline used for the Chinese data to obtain daily measures of U.S. war-related diplomatic sentiment.

The Chinese and U.S. corpora are complementary and institutionally distinct. U.S. releases more often address conflicts and security issues involving third countries, including civil wars, interventions, terrorist incidents, and peacekeeping operations, consistent with the global orientation of U.S. security policy. Chinese press conferences focus more heavily on China's own security environment, including regional disputes, Taiwan-related sovereignty issues, border clashes, and military activity in East Asia. The Chinese corpus is therefore weighted towards East Asian and near-abroad tensions, whereas the U.S. corpus contains more statements on conflicts in the Middle East, Africa, and Europe, alongside commentary on China-U.S. military frictions. For example, in an 8 July 2025 press statement, the Department Spokesperson "condemn[ed] the unprovoked Houthi terror attack" on a civilian cargo ship in the Red Sea, using strong negative language to denounce the perpetrators and express solidarity with affected parties. Such releases provide clear instances of strongly negative war-related sentiment analogous to the Chinese example above.

Taken together, the Chinese MFA press conferences and U.S. Department of State official releases illustrate why official diplomatic texts are useful for high-frequency sentiment measurement. Because the texts come from central government communications, the WDSI is anchored in deliberate policy rhetoric and avoids media framing or ex post commentary. These two corpora span more than fifteen years of high-frequency observations and provide the two most detailed country examples within the broader 15-economy WDSI database.

III. Detailed Procedure for Constructing the WDSI

This section describes the technical procedure used to transform official diplomatic texts into country-level war-related diplomatic sentiment indices. We use the Chinese Ministry of Foreign Affairs press-conference corpus as the detailed implementation example because it is the main daily series in the baseline flow analysis and has a transparent Q&A structure. The same multi-stage scoring architecture is applied to the U.S. corpus and to the other country corpora in the 15-economy database summarised in Online Appendix II, but the text unit supplied to the model depends on the structure of the source.

A. Data Pre-processing and Input Units

The cross-country implementation starts from source-specific text records. For Chinese MFA press conferences and other explicitly dialogic briefings, the natural input unit is a question-answer (Q&A) exchange. For stand-alone releases, statements, readouts, remarks, speeches, and communiques, the input unit is the self-contained source record; when a release page contains clearly separable sections or multiple independent items, the text is segmented at those source-provided boundaries. Non-Q&A corpora retain their source-provided structure. This design preserves the institutional form of each country's official communication while keeping the downstream scoring task the same: each input passage is screened for war/security relevance and, if relevant, assigned a direction and intensity score.

The Chinese MFA corpus covers 5 January 2010 through 17 April 2026. We collect the full regular press-conference transcripts over this window. Each transcript is divided into Q&A units, where a unit consists of one journalist question and the corresponding spokesperson reply. We remove boilerplate text, including greetings, housekeeping remarks, and logistical announcements, and discard non-substantive fragments such as incomplete sentences or purely procedural exchanges. The cleaned Q&A units are scored at the unit level and then aggregated to the calendar-day level. For non-Q&A corpora, cleaned publication records are scored as source-specific passages and then aggregated by country and publication date. Days without a new publication remain missing until the index-construction step, where the most recent observed diplomatic stance is carried forward to align the WDSI with daily financial-market observations.

B. Three-stage LLM Pipeline

Sentiment extraction uses a vertically decomposed three-stage pipeline. The design separates military relevance, sentiment direction, and sentiment intensity into distinct tasks, with one LLM "agent" assigned to each step:

(i) Military relevance screening (Agent 1). Agent 1 reads each source-specific text passage and determines whether it concerns war, armed conflict, the use of force, military security, or closely related escalation risks. It returns a binary war-related label and a concise natural-language rationale. Only passages that pass this screen proceed to the sentiment-direction and intensity stages; all other passages contribute zero war-related sentiment to the daily index.

(ii) Sentiment category classification (Agent 2). For war-related passages, Agent 2 classifies the direction of the statement with respect to war, conflict, and the use of force as negative, neutral, or positive. Negative sentiment covers condemnation, explicit or implicit threats, strong warnings, and other language that signals escalation or opposition. Positive sentiment covers support for de-escalation, conflict resolution, humanitarian relief, or peaceful outcomes. Neutral sentiment is reserved for factual or balanced statements that contain no clear evaluative stance. Agent 2 reports the category and a short rationale identifying the phrases that drive the classification.

(iii) Intensity scoring (Agent 3). Conditional on the Stage (ii) direction, Agent 3 maps the passage to a discrete intensity score on the $[-3, 3]$ scale. Negative statements receive -1, -2, or -3, corresponding to mild, moderate, or strong negative sentiment; positive statements receive $+1$, $+2$, or $+3$ on the analogous positive scale. Neutral statements are coded as 0. The prompt instructs Agent 3 to distinguish mild concern from firm opposition and explicit threats of countermeasures, and to justify the assigned score in a short explanation.

C. *Prompt Design and Implementation*

All three agents use the same underlying LLM but receive task-specific prompts. Each prompt defines the task in plain language, lists the admissible labels or scores, and requests both a machine-readable output and a concise rationale. The Stage 2 and Stage 3 prompts explicitly condition on the previous-stage output, which prevents mechanically inconsistent combinations of direction and intensity. This decomposition shortens the reasoning chain at each step and makes the measurement auditable because every label or score is paired with an explanation.

D. *Validation and Error Control*

To reduce random classification error, each source-specific passage is evaluated three times at each stage using prompt variants or different random seeds. A separate validation agent then compares the independently generated labels, scores, and rationales. (i) If the runs are consistent, the common output is accepted. (ii) If discrepancies arise, the validation agent examines the underlying rationales, selects the most coherent interpretation, or directs the pipeline to re-run the stage using a more precise prompt.

This consensus procedure reduces random misclassifications and is particularly valuable in edge cases, such as indirect references to military issues, rhetorical questions, coded language, and multi-country comparisons. Manual audits of the stratified subsample from the Chinese MFA, detailed in Online Appendix IV, indicate that the pipeline achieves over 85% accuracy at each individual stage and exceeds 75% accuracy for the full three-stage classification. These results support the use of the LLM-based measure as the primary input for WDSI, while the robustness checks benchmark it against simpler text-classification alternatives.

E. *From Scores to WDSI*

Each scored source-specific passage receives one raw sentiment score in $\{-3, -2, -1, 0, 1, 2, 3\}$, with lower values indicating more negative war-related diplomatic sentiment. For publication days with more than one scored passage, we aggregate to the country-day level by retaining the most adverse war-related signal observed on that day, that is, the minimum score across same-day passages for each sentiment category. This rule gives priority to the strongest official war-related signal released on a given day and avoids averaging away short but salient negative statements. Source-specific passages with no war-related content receive a score of 0. Non-publication days are forward-filled with the most recent observed country-day score, reflecting the persistence of official diplomatic positions between releases. The reported WDSI is the 7-day rolling average of this daily score, which smooths idiosyncratic single-publication noise while preserving high-frequency variation.

IV. Validation and Robustness of the WDSI

A. Manual Annotation Validation

We begin by benchmarking the pipeline against human-coded labels. From the full Chinese MFA corpus, we draw a stratified random sample of approximately 1,200 Q&A units at roughly equal monthly intervals over January 2010 to April 2026. Two research assistants with training in international relations independently annotate each selected unit.

Each annotator, with no access to the model outputs, records three labels: whether the Q&A unit concerns war or military security; the sentiment category, coded as negative, neutral, or positive for war-related units; and the sentiment intensity. Inter-annotator agreement is about 88% for the binary war-related indicator, 80% for sentiment direction, and 83% for sentiment intensity. We use the reconciled human labels as the validation benchmark. Relative to this benchmark, the multi-stage pipeline exceeds 85% agreement at each individual stage, and the share of Q&A units for which the full three-stage extraction is correct exceeds 75%. The remaining errors are concentrated in borderline or ambiguous cases, including indirect allusions to conflict and mixed diplomatic messages, which are also the cases in which human coders disagree most often.

The validation sample also provides little evidence of directional error. The pipeline does not systematically overstate negativity, nor does it regularly classify economic disputes as military events. This pattern is important because either error would mechanically bias the WDSI downwards or contaminate it with non-military diplomatic rhetoric.

B. Model Comparison and Selection

We next use the human-labelled sample to compare the LLM pipeline with two lower-dimensional text-analysis benchmarks: a dictionary classifier and a shallow supervised classifier.

The first benchmark is a dictionary classifier built from curated military terms and sentiment-laden keywords. It flags war-related content using terms such as "war", "military", and "attack", and measures polarity by counting negative and positive phrases, such as "condemn" or "warn" versus "welcome" or "commend". The second benchmark is a shallow supervised classifier: we train a logistic model on pre-trained Word2Vec features to predict war relevance and sentiment using the manually labelled subset. We then compare these benchmarks with several LLMs capable of processing multilingual diplomatic text. We adopt *Qwen-3* in the main pipeline because it performs best on the validation tasks reported below.

Table IV.1 reports the validation results. Both low-dimensional benchmarks fall well short of the LLM-based methods. The dictionary classifier is especially weak on sentiment direction and intensity, with overall accuracy in the 50–60% range. The Word2Vec-logit benchmark improves on this baseline but reaches only about two-thirds accuracy. Qwen-3 already dominates the non-LLM alternatives in a one-step specification. The full multi-step pipeline performs best, with war-relevance accuracy above 90% and full-classification accuracy close to 80%. The evidence supports using the decomposed LLM pipeline as the main input to the WDSI because it recovers diplomatic nuance that coarser text methods miss.

Table IV.1: Classification Accuracy of Different Sentiment Extraction Methods

The table reports the classification accuracy of various methods at key stages of the sentiment extraction task, evaluated on a human-labelled sample of approximately 1,200 Chinese MFA Q&A units. *War detection* refers to correctly identifying whether a unit is war/military-related. *Sentiment category* accuracy is the percentage of war-related units for which the method correctly classifies the sentiment as negative, neutral, or positive. *Sentiment intensity* accuracy is the percentage of war-related units where the strength of sentiment (on the -3 to +3 scale) is correctly assigned. *Overall* accuracy is the share of all units (war-related or not) for which the method's complete output matches the ground-truth label (including correctly outputting "no war content" for non-war units). The LLM methods use the Qwen-3 model. The multi-step pipeline corresponds to our full procedure with staged prompts and consensus checks, whereas the direct single-step LLM approach attempts to produce the sentiment score in one shot. The LLM pipeline exhibits the highest accuracy at each stage and overall.

Method	War detection (%)	Sentiment category (%)	Sentiment intensity (%)	Overall (%)
Dictionary-based approach	81.4	65.0	57.2	52.3
Shallow ML model (embeddings + LR)	86.5	74.1	70.3	64.1
LLM (direct single-step)	90.2	82.0	75.6	72.5
LLM (multi-step pipeline)	93.3	88.4	87.1	79.8

C. *Single-step versus Multi-step Strategy*

The validation exercise also shows why the WDSI uses a multi-step LLM design. In the "direct scoring" strategy, the model assigns the final sentiment score, or a "no-war" designation, in a single prompt. This approach is faster but less reliable: it sometimes misses implicit military content and sometimes treats mildly negative diplomatic language as neutral. Table IV.1 shows that its overall accuracy is about 5–10 percentage points below that of the structured pipeline. The multi-step procedure lowers this error rate by separating topical relevance, sentiment direction, and sentiment intensity. It also adds a consensus check. In the Chinese MFA validation sample, each agent is run three times per Q&A unit using slight prompt variations; when outputs disagree, a validation routine reviews the rationales and, if needed, re-queries the model with a clarified prompt. In the validation sample, about 12% of war-related units initially produce inconsistent outputs at one stage. Most are resolved by a single automated retry, and fewer than 3% require a second retry or manual inspection. A parallel exercise with human coders points in the same direction: inter-coder agreement falls by roughly 10 percentage points when coders assign an overall sentiment score directly without following the stepwise protocol. The evidence therefore favors vertical decomposition because it improves both accuracy and interpretability.

D. *Robustness Checks*

We next verify that the WDSI is robust to specific implementation choices. First, we vary the smoothing and missing-value rules. The main index uses a 7-day rolling average and forward-fills non-conference days. Using 5- or 10-day windows produces nearly identical series: pairwise correlations with the baseline exceed 0.95. Alternative imputations, including backward filling or leaving gaps and interpolating monthly means, also preserve the main trends and turning points. Second, we test model dependence by re-running the pipeline on a subset of the data with GPT-4, accessed through the OpenAI API, as an alternative to Qwen-3. The alternative classifications are strongly correlated with the main outputs, with Pearson $r \approx 0.8$, and show no systematic shift in average sentiment. We retain Qwen-3 because it performs slightly better in the validation exercise and handles multilingual diplomatic text directly; the substantive signal is robust across model choices. Third, we examine error patterns for directional bias. We find no evidence that the pipeline systematically overstates negativity. False negatives in war detection are concentrated in genuinely subtle cases, such as indirect references to security alliances, while false positives are rare. Sentiment-classification errors are split mainly between neutral versus mild negative language and mild versus moderate negativity. These patterns support the interpretation of the WDSI as a stable measure of official war-related diplomatic sentiment, with little dependence on a particular smoothing rule, imputation method, or model choice.

E. Supplementary Inference Checks for Baseline Flow Regressions

Table IV.2 assesses the sensitivity of the baseline China flow results in Table 4 to alternative inference procedures. We re-estimate the same specifications and report the coefficient on $WDSI_t^{CN}$ using Newey–West HAC standard errors and moving-block bootstrap inference, with 5- and 22-trading-day bandwidths or block lengths.

F. Supplementary Granger-Causality Check: A-H Premium and Stock Connect Flows

As a final mechanism check, we estimate residualised daily VAR(p) systems linking the A-H share premium to each Stock Connect flow series. Before estimation, both variables are residualised on the full baseline control set and month-of-year fixed effects, so the tests ask whether remaining variation in relative prices predicts remaining variation in flows. Table IV.3 reports Granger-causality p-values for lag orders $p = 1, 3, 5$. The strongest and most stable pattern is A-H premium $\rightarrow$ outflow, which holds across all three lag choices. Net-flow predictability appears at shorter horizons, while the inflow evidence is weaker and mainly reflects reverse predictability from inflows to the premium. The pattern is consistent with the proposed mechanism: sentiment-induced relative-price movements precede the southbound flow response.

Table IV.2: Supplementary Inference Checks: HAC and Moving-Block Bootstrap

This table re-estimates the baseline China regressions in Table 4 and reports the coefficient on the Chinese war-related diplomatic sentiment index $WDSI_t^{CN}$ under alternative inference methods. Panels A-B use Newey–West HAC standard errors with lag lengths 5 and 22. Panels C-D use moving-block bootstrap (MBB) with block lengths 5 and 22, respectively, based on 999 bootstrap replications; bootstrap standard errors are the standard deviation of bootstrap coefficients. For the bootstrap panels, significance stars use the empirical bootstrap p-value, so star assignments can differ from coefficient-to-standard-error ratios under a normal approximation. All specifications include the same control block and month-of-year fixed effects as the baseline model. Standard errors are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) Inflow	(2) Outflow	(3) Net flow
Panel A. Newey–West HAC			
$WDSI^{CN}$	0.009	-0.299***	0.424*
	(0.161)	(0.106)	(0.217)
Lag / block length	5	5	5
Controls	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes
Observations	2,205	2,224	2,190
Panel B. Newey–West HAC			
$WDSI^{CN}$	0.009	-0.299**	0.424*
	(0.159)	(0.140)	(0.255)
Lag / block length	22	22	22
Controls	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes
Observations	2,205	2,224	2,190
Panel C. Moving-block bootstrap			
$WDSI^{CN}$	0.009	-0.299***	0.424*
	(0.166)	(0.105)	(0.213)
Lag / block length	5	5	5
Controls	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes
Observations	2,205	2,224	2,190
Panel D. Moving-block bootstrap			
$WDSI^{CN}$	0.009	-0.299**	0.424
	(0.172)	(0.140)	(0.268)
Lag / block length	22	22	22
Controls	Yes	Yes	Yes
Month-of-year FE	Yes	Yes	Yes
Observations	2,205	2,224	2,190

Table IV.3: Supplementary Granger-Causality Results: A-H Premium and Stock Connect Flows

This table reports Granger-causality tests from residualised daily VAR(p) models between the A-H premium and Stock Connect flows (Inflow, Outflow, Net flow), for lag orders $p = 1, 3, 5$. Both series are residualised using the same baseline control block and month-of-year fixed effects as Equation (1). "p-value: A-H→Flow" tests whether lagged A-H premium predicts the flow variable; "p-value: Flow→A-H" reports the reverse-direction test.

Flow channel	Lag order	p-value: A-H → Flow	p-value: Flow → A-H	Direction at 5%
Inflow	1	0.0489	0.000000844	Bidirectional
Inflow	3	0.3991	0.00000121	Flow → A-H
Inflow	5	0.7397	0.00000997	Flow → A-H
Outflow	1	0.0000000259	0.7791	A-H → Flow
Outflow	3	0.0000484	0.4529	A-H → Flow
Outflow	5	0.000468	0.5291	A-H → Flow
Net flow	1	0.0000718	0.000791	Bidirectional
Net flow	3	0.0311	0.000429	Bidirectional
Net flow	5	0.1621	0.00244	Flow → A-H