國立臺灣大學社會科學院經濟學系

碩士論文

Department of Economics

College of Social Sciences

National Taiwan University

Master's Thesis

提升地緣政治風險評估：以台灣爲應用案例的創新大型語言模型方法

Improving Geopolitical Risk Assessment: A Novel LLM-Based Approach with Application to Taiwan

何主恩

Chu-En Ho

指導教授：林明仁 博士、陳由常 博士

Advisors: Ming-Jen Lin, Ph.D. and Yu-Chang Chen, Ph.D.

中華民國 114 年 7 月

July, 2025

doi:doi:10.6342/NTU202503488

摘要

傳統以固定、手工編製字典爲基礎的地緣政治風險（GPR）指數，往往忽略在地脈絡與混合戰手法的演變。本論文提出並驗證一套由大型語言模型（LLM）輔助的「提示—篩選—計數」工作流程：首先生成 40 份以臺灣爲焦點的華語 GPR 字典，將每份字典轉換爲週頻指數，再取其中位數作爲主指標。依據 2017─2024 年共 330 萬則新聞報導，這種指數組合設計可降低 LLM 輸出隨機性（平均配對相關係數 = 0.70），並維持可解釋性。相較於 Caldara 的全球字典與 Lau 的臺灣改編版本，我們的指數涵蓋更廣泛的風險面向—軍事升級、外交孤立、經濟脅迫、網路／資訊攻擊與國內政治。事件研究顯示，本指數能捕捉傳統指數遺漏的臺灣特定衝擊：經濟排斥（臺灣被排除於 IPEF）、北京主導的邦交國轉向（宏都拉斯與諾魯）、區域安全震盪（南韓 2024 戒嚴風波），以及領導人更迭不確定性（習近平三連任與 2024 年美國總統大選）。來源層級分析發現《自由時報》波動劇烈，《中國時報》輸出較穩定卻呈現內容農場現象，提醒研究者在計算新聞占比型指數時須慎防分母偏誤。七變數季頻結構向量自迴歸（SVAR）結果顯示，臺灣經濟韌性十足：一個兩倍標準差的 GPR 衝擊後，投資增加、工時延長、股價上升，而隱含波動率僅短暫下跌。此方法成本低、人工校對需求少、可解釋度高，可輕鬆擴充至其他議題，例如跨國或台積電專屬的風險指數。

關鍵字：大型語言模型、地緣政治風險、文本分析

doi:doi:10.6342/NTU202503488

doi:doi:10.6342/NTU202503488

Abstract

Geopolitical risk (GPR) indices based on fixed, handcrafted dictionaries often miss local context and evolving hybrid-warfare tactics. This thesis introduces an LLM-assisted "prompt—filter—count" workflow that generates 40 Mandarin, Taiwan-specific GPR dictionaries, converts each into a weekly series, and takes their median as the headline index. Using 3.3 million news articles (2017—2024), the ensemble mitigates LLM output randomness (mean pair-wise correlation = 0.70) while preserving transparency. The median index captures a wider range of risks—military escalation, diplomatic isolation, economic coercion, cyber/information attacks, and internal politics—than Caldara's global list and is more balanced than Lau's Taiwan adaptation. Event-study tests show the index flags Taiwan-specific shocks that legacy lists miss: economic exclusion (Taiwan's IPEF omission), Beijing-driven ally switching (Honduras & Nauru), regional-security shake-ups (South Korea's 2024 martial-law scare), and leadership-transition uncertainty around Xi's third term and the 2024 U.S. election. Source-level analysis reveals sharp contrasts: Lib-

doi:doi:10.6342/NTU202503488

erty Times shows episodic surges, China Times steadier output but signs of content farming, cautioning against naïve denominator use in news-share indices. A seven-variable quarterly SVAR indicates Taiwan's economy is resilient: a two-s.d. GPR shock precedes higher investment, longer work hours, rising equity prices, and a short-lived dip in implied volatility. The method's low cost, minimal curation, and interpretability make it adaptable to other contexts, such as TSMC-specific or cross-country risk indices.

Keywords: large language model, geopolitical risks, text analysis

doi:doi:10.6342/NTU202503488

Contents

doi:doi:10.6342/NTU202503488

doi:doi:10.6342/NTU202503488

List of Figures

doi:doi:10.6342/NTU202503488

doi:doi:10.6342/NTU202503488

List of Tables

doi:doi:10.6342/NTU202503488

doi:doi:10.6342/NTU202503488

Chapter 1 Introduction

Geopolitical risk (GPR) has moved from the periphery to center stage, even more visibly after Russia's invasion of Ukraine in 2022. Building on the computational gains and large-scale news digitization of the last decades, Caldara and Iacoviello [2018, 2022] proposed a simple yet powerful idea: construct a dictionary of GPR related terms, count their frequency in major English language outlets, and normalise that share into an index. Their global GPR series is now a "workhorse" variable in studies that range from the oil-stock nexus over 1899—2016 [Antonakakis et al., 2017], to volatility spill-overs [Smales, 2021], quantile responses in BRIC equities [Rawat and Arif, 2018], and the macro consequences of shock size [Brignone et al., 2024] and inflation[Caldara et al., 2022]. Yet two intuitive mismatches stand out. First, the Caldara—Iacoviello approach is U.S. centric and English centric; translating it directly into Mandarin news may miss Taiwan specific risk cues. Second, Taiwan's primary external threat is China, whose toolkit increasingly relies on hybrid and cognitive warfare such as diplomatic isolation, disinformation, and economic coercion [Hung and Hung, 2022, Pietrzak, 2025]. If warfare itself evolves, the words that proxy for that warfare must evolve as well.

Creating such an updated dictionary by hand is daunting: hybrid operations cut across the military, political, economic, cyber and informational spheres. Instead, this thesis asks: Can large language models (LLMs) generate a Taiwan tailored, Mandarin GPR dictionary

doi:doi:10.6342/NTU202503488

that outperforms general purpose manually crafted dictionary?

We choose to keep with the dictionary approach rather than a full LLM classification pipeline. First, ask LLM to generate a Taiwan-specific dictionary. Second, follow the traditional Caldara frequency count mechanism with the dictionary, yielding the GPR index. Third, rather than trusting a single LLM output, we prompt four frontier models across five institutional "roles," run each prompt ten times, retain 40 valid dictionaries, and compute 40 weekly indices. Taking their median yields a stable Taiwan-GPR series that is robust to prompt stochasticity. This two-step design keeps token costs manageable, preserves interpretability, and allows analysis of why a particular word inflates or deflates the index, which is a crucial feature given earlier evidence of LLM bias and context blindness[Li et al., 2024].

The thesis is structured as follows.

1. **Data and dictionary generation** (chapter 2). We use 3.3 million Mandarin articles (2017—2024) from three leading Taiwanese outlets and prompt GPT-4.1 to propose Taiwan specific GPR dictionary.

2. **Index construction and diagnostics** (section 3.1). We compare the new series against the Taiwanese adaptation by Lau [2024].

3. **Spike attribution and category analysis** (section 3.2 - section 3.4). Event studies reveal that LLM generated word lists tied to diplomatic crisis and internal political instability flag additional peaks that the traditional dictionary misses.

4. **Macroeconomic relevance** (section 3.5). A structural VAR model shows that Taiwan is resilient to geopolitical risk increase.

doi:doi:10.6342/NTU202503488

Our contributions are twofold: on one hand, we introduce an LLM-assisted dictionary workflow that balances scalability with transparency. On the other hand, we deliver the first Mandarin GPR index that accounts for hybrid warfare signaling of Taiwan.

The remainder of the thesis is organized as outlined above and concludes in chapter 4 with limitations and avenues for future research.

doi:doi:10.6342/NTU202503488

doi:doi:10.6342/NTU202503488

Chapter 2 Data and Dictionary Construction

2.1 Data Sources

We use CommonCrawl News(CC-News) as our news source. CC-News is a dataset that contains news articles from all over the world and is released by Common Crawl. Common Crawl continuously releases freshly crawled data on a sub-day basis, which is different from the Common Crawl corpus (CC-MAIN) monthly release. The dataset started in late 2016 to this day. It is web crawl data, and suffers from the limitations of internet data. Websites' revisions and policy changes, crawling list changes, network issues, etc. all may affect the crawled data.

We filtered the dataset to only leave three outlets: China Times (中國時報), United Daily (聯合報), Liberty Times (自由時報). One of the reasons to choose them is that they are the only three privately owned, general-interest Chinese-language newspapers in Taiwan, while Lau [2024] also chose China Times and United Daily. They also represent different political positions, covering both pro-KMT and pro-DPP. Another reason is that they have more continuous data in history, which other outlets were missing over one-third of time periods. We further limited the date range to 2017-01-01 to 2024-12-

doi:doi:10.6342/NTU202503488

31, because both the start and end of data have heavy missing issues. The final dataset includes 3,315,530 articles, consisting of 45% China Times, 38.5% United Daily, and 16.4% Liberty Times.

2.1.1 Missing Data Patterns

Despite selecting outlets with better coverage, missing data remains present in our dataset. China Times has 68 missing weeks out of 418 (16.3%), Liberty Times has 44 missing weeks (10.5%), and United Daily has 74 missing weeks (17.7%). The missing patterns reveal two distinct causes:

First, the substantial missing data in 2017, particularly for United Daily (47 consecutive weeks), reflects the early instability of CC-NEWS's site list. United Daily appears to have been added to the crawling list only after 2018.

Second, the increasing missing data from late 2023 onwards across all newspapers likely results from websites implementing bot-blocking measures in response to increased LLM-based crawling. We confirmed that China Times currently blocks CCBot in their robots.txt file, though the exact timing when this blocking began is unknown.

These missing data patterns are important limitations to consider when interpreting GPR index results, particularly for 2017 (where United Daily's perspective is largely absent) and 2024 (where all three newspapers show significant gaps).

doi:doi:10.6342/NTU202503488

2.2 LLM Dictionary Generation: A Two-Step Approach

We employ a two-step LLM prompting strategy to generate Taiwan-specific GPR dictionaries. This approach addresses LLM output instability while enabling structured analysis through fixed categories.

2.2.1 Step 1: Role-Based Dictionary Generation

The first step explores diverse institutional perspectives on Taiwan's geopolitical risks. We prompt four models (Claude-Opus-4, Gemini-2.5-Pro, GPT-4.1, GPT-4o) to assume five distinct roles: president, defense minister, National Security Bureau (NSB) director, political science professor, and journalist. We attempted 40 dictionary generations (4 models × 5 roles × 2 runs), but 5 failed validation by using forbidden OR operators, yielding 35 valid dictionaries.

The prompt structure remains minimal:

1. Role specification (e.g., "You are Taiwan's president")
2. Task definition ("Generate a GPR dictionary for Taiwan")
3. Output format (categories with AND-operator word sets)

The complete prompt templates for both Step 1 and Step 2 are provided in Appendix section A.1 and Appendix section A.2.

Step 1 reveals striking similarity across roles: all identify military threats, diplomatic pressure, economic coercion, and information warfare as core risks. The 35 dictionaries

doi:doi:10.6342/NTU202503488

collectively generate 172 unique category names, but these largely represent naming variations of the same concepts. For instance, “Military Escalation”, “Military_Tensions”, “軍事威脅”, and “direct_military_threat” all capture similar risk dimensions.

However, the resulting GPR indices show low pairwise correlations (see section 3.1), suggesting substantial variation in word selection despite conceptual similarity. This motivates Step 2: rather than choosing a single role’s perspective, we seek to combine insights from all five institutional viewpoints into a unified framework.

2.2.2 Step 2: Fixed-Category Consolidation

The second step standardizes outputs through fixed-category prompting. We analyzed 172 unique categories from Step 1 and consolidated them into eight representative groups based on conceptual overlap and coverage completeness:

- Military Escalation
- Diplomatic Isolation
- Economic Coercion
- Regional Security
- Internal Political Instability
- Gray Zone Operations
- US-China-Taiwan Relations
- Cyber and Information Warfare

These eight categories capture almost all risk dimensions from step 1 results. Notably, we merged ’cross-strait political tensions’ into Military Escalation since political

doi:doi:10.6342/NTU202503488

rhetoric without military backing rarely constitutes significant geopolitical risk in Taiwan's context.

For Step 2, we instruct the same four models to generate word sets for each fixed category. Each model runs 10 times, producing 40 dictionaries with an identical structure. We calculate the median GPR index across dictionaries with confidence intervals, addressing the single-dictionary limitation of traditional approaches.

2.3 Result Dictionary

Table 2.1 presents a representative Step2 dictionary showing our fixed eight-category structure. Compared to Caldara's dictionary, there are several differences that need to be clarified. First, we discard proximity search, but LLM yields more specific and long terms which balance the question that it may not capture relevant risks. For example, 「美國介入」can be seen as two consecutive words, 「美國」and 「介入」which is a stronger version of proximity search. Second, unlike English where words may appear in multiple forms (e.g., plurals or past tense) that need to be included in the dictionary as Caldara did, Chinese words usually remain unchanged across such variations, simplifying the dictionary. Third, we did not try to choose core words, nor try to distinguish acts and threats categories, which contradicts our goal to conduct GPR analysis with lacking domain knowledge.

Our dictionary also captures a much wider set of hybrid warfare concepts than the manual dictionaries such as Lau [2024] and Caldara and Iacoviello [2022]. Seven of the eight categories do not belong to traditional warfare risks. This difference indicates that the model reflects local and recent trends. Taiwan's geopolitical main concern is to deter

doi:doi:10.6342/NTU202503488

a possible invasion from China, while China seeks to counter such preparations. Thus, related terms such as economic coercion, potential foreign intervention, diplomatic isolation, infrastructure threats, and cyber and information attacks are relevant. Words describing direct war acts are less frequent in the dictionary probably because no war act has occurred in decades around Taiwan. The same reason makes a standard global dictionary such as Caldara and Iacoviello [2022] less suitable for Taiwan, resulting in Lau's GPR index containing one dominant category contributing 86% of the index since those direct act categories are nearly impossible to be mentioned in Taiwan's news. On the contrary, the LLM yields a more balanced set of categories, with the top contributor accounting for only 42%. This suggests our index captures a broader range of risks.

Consequently, the LLM offers a more context-aware and accurate dictionary on Taiwan-specific geopolitical risk.

doi:doi:10.6342/NTU202503488

Table 2.1: Example Search Queries and Word Sets[1]

Panel A. Search categories and search queries

Category	Search query
Military Escalation	military_threats AND cross_strait_tensions
Internal Political Instability	domestic_unrest AND political_polarization
Diplomatic Isolation	diplomatic_pressure AND international_recognition
Regional Security	regional_tensions AND alliance_security
Economic Coercion	economic_sanctions AND trade_warfare
Gray Zone Operations	hybrid_threats AND gray_zone_tactics
US-China-Taiwan Relations	us_china_competition AND taiwan_strait_dynamics
Cyber & Information Warfare	cyber_attacks AND cognitive_warfare

Panel B. Search word sets (selected examples)

Word set	Keywords / phrases
military_threats	軍事威脅，武力犯台，軍事衝突，戰爭風險，武力統一，軍事挑釁，飛彈威脅，軍事恫嚇，武裝衝突，軍事對峙，戰爭危機，軍事入侵，武力威脅，軍事封鎖，戰爭陰影，軍事冒險，武力脅迫，軍事緊張，戰爭邊緣
cross_strait_tensions	兩岸緊張，台海危機，海峽中線，共機擾台，軍演威嚇，圍台軍演，越過中線，防空識別區，軍機繞台，軍艦擾台，實彈演習，聯合軍演，封鎖演習，導彈試射，兩岸對峙，軍事演訓，跨越海峽，台海局勢，兩岸軍事
diplomatic_pressure	外交施壓，國際孤立，外交打壓，斷交危機，邦交流失，外交圍堵，國際封殺，外交戰，邦交國，外交困境，國際打壓，外交挫敗，斷交威脅，外交孤立，國際排擠，外交危機，邦交動搖，外交封鎖
economic_sanctions	經濟制裁，貿易制裁，金融制裁，經濟封鎖，出口管制，投資限制，技術封鎖，經濟懲罰，貿易限制，金融脫鉤，經濟孤立，制裁威脅，經濟報復，貿易壁壘，經濟圍堵，產業制裁，科技制裁，經濟打壓
cyber_attacks	網路攻擊，駭客入侵，資安威脅，網路戰，惡意軟體，資料外洩，網路癱瘓，駭客攻擊，資安漏洞，網路入侵，系統攻擊，網路間諜，資安事件，網路破壞，駭客威脅，資訊竊取，網路滲透，系統癱瘓
regional_tensions	區域緊張，印太戰略，南海爭議，東海緊張，區域衝突，亞太局勢，區域對抗，地緣緊張，區域不穩，印太局勢，區域危機，亞太緊張，區域對峙，地緣衝突，區域動盪，印太對抗，區域安全，亞太危機
us_china_competition	美中對抗，大國競爭，美中角力，戰略競爭，美中衝突，新冷戰，大國博弈，美中脫鉤，戰略對抗，美中緊張，霸權競爭，美中摩擦，地緣競爭，美中較量，戰略角力，大國對峙，美中爭霸，戰略競逐
gray_zone_tactics	灰色地帶，海警船，民兵船，漁船包圍，海上民兵，準軍事，灰色手段，邊緣策略，灰色戰術，非戰爭，準軍事行動，灰色作戰，邊緣行動，灰色衝突，海上灰色，準戰爭行爲，灰色挑釁，邊緣政策

doi:doi:10.6342/NTU202503488

doi:doi:10.6342/NTU202503488

Chapter 3 Empirical Results

3.1 GPR Index: Computation & Time-Series Results

To build the GPR index, articles are flagged as GPR relevant if they satisfy at least one of the category queries. Each category represents a specific channel through which geopolitical risk manifests and is formalized as a Boolean search query combining word sets from the dictionaries described in Table 2.1.

Aggregation is performed weekly to average out day-of-week effects. The index for week t is computed as:

$$\text{GPR Index}_t = \frac{\#\,\text{flagged articles}_t}{\text{Total articles}_t}. \tag{3.1}$$

This value represents the share of total news coverage that signals potential geopolitical risk. We then scale each index to mean = 100 following Caldara and Iacoviello [2022].

Using the 40 Step2 dictionaries generated through our fixed 8-category approach, we calculate 40 individual GPR indices and take their median as our primary measure. This ensemble approach addresses LLM randomness concerns while ensuring all dictionaries

doi:doi:10.6342/NTU202503488

share the same category structure. The mean pairwise correlation among the 40 Step2 indices is 0.698, indicating substantial agreement. In contrast, the mean pairwise correlation among the 35 Step1 indices is 0.493, demonstrating that the fixed-category approach in Step2 provides notably higher consistency.

For comparison, we also compute Lau [2024]'s GPR index. The correlation between our Step2 median index and Lau's index is 0.611, demonstrating strong alignment despite different methodological approaches.

Figure 3.1 presents the GPR index, while Figure 3.3 provides a direct comparison between our index and Lau's index on the same scale. The peaks in Figure 3.2 are identified by any continuous weeks that indexes are larger than one standard deviation above the mean. Most trends and peaks are similar, but both indices have some unique peaks that only appear in one version. We discuss those distinct peaks that only appear in our version in the next section (section 3.2).

We also examine the trend and levels of each category in Figure 3.4. For each category, we calculate 40 step 2 separate indices using that category's article counts as the numerator and total articles as the denominator. We then take the median across these 40 indices for each week, following the same ensemble approach used for the overall GPR index. The results show each category's GPR index having peaks at different times confirms that categories are moderately balanced and capture distinct risk dimensions even though no balance requirement was specified in the prompt.

[0]We use only Lau's dictionary to calculate this index. The data sources and index construction methodology follow Caldara's approach, not Lau's method.

doi:doi:10.6342/NTU202503488

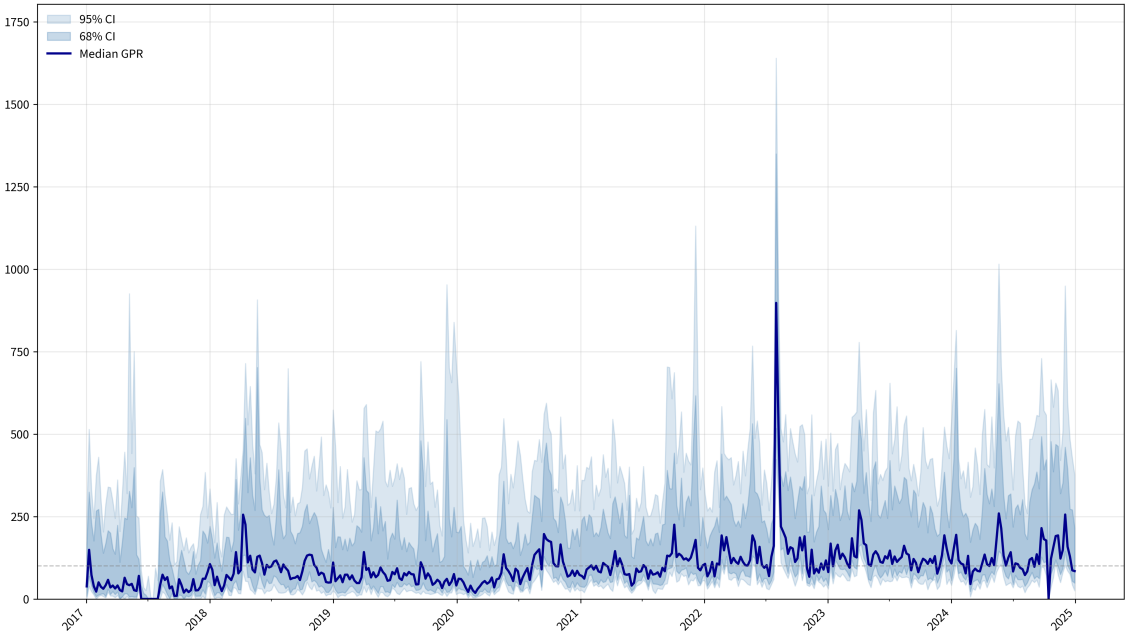


Figure 3.1: Weekly GPR Index (Median and band of 40 Step2 indices)

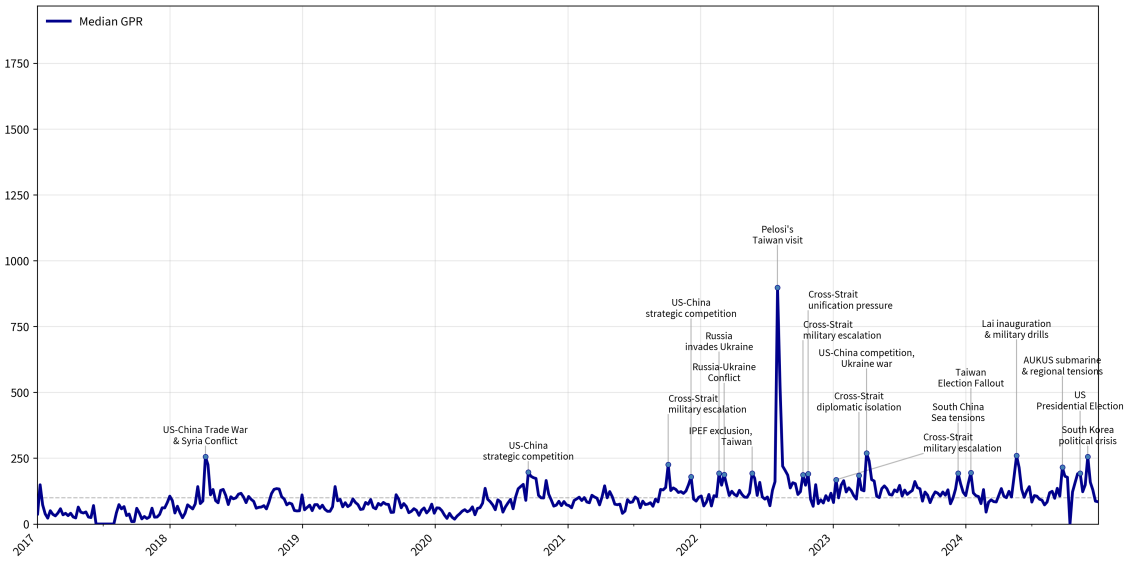


Figure 3.2: Weekly GPR Index with spikes tagged

doi:doi:10.6342/NTU202503488

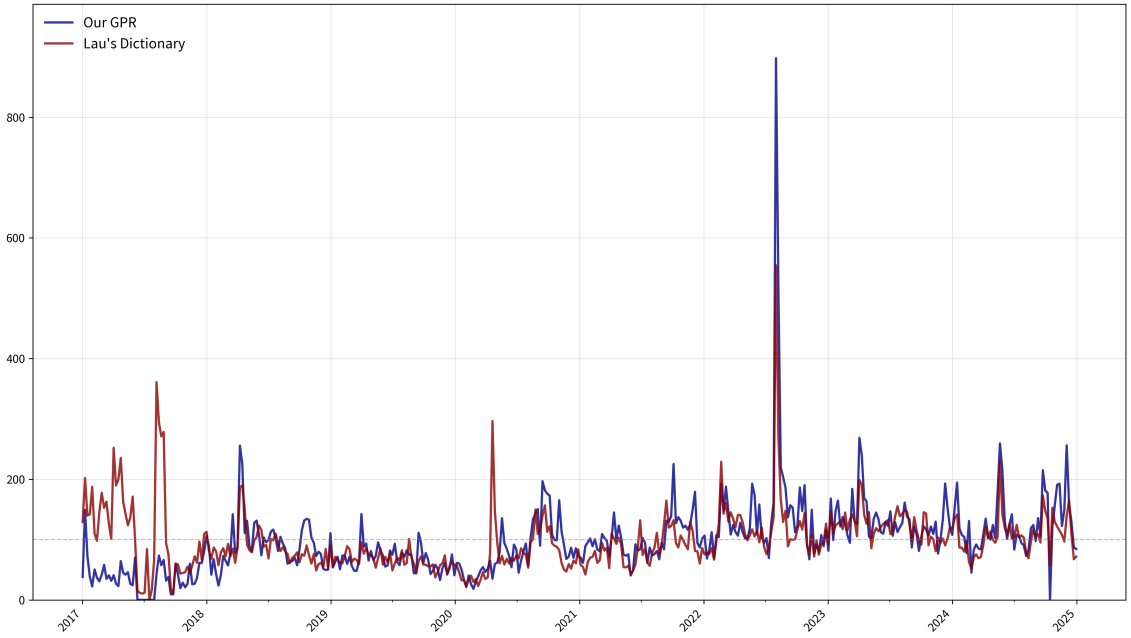


Figure 3.3: Weekly GPR Index: Compare with Lau's Dictionary

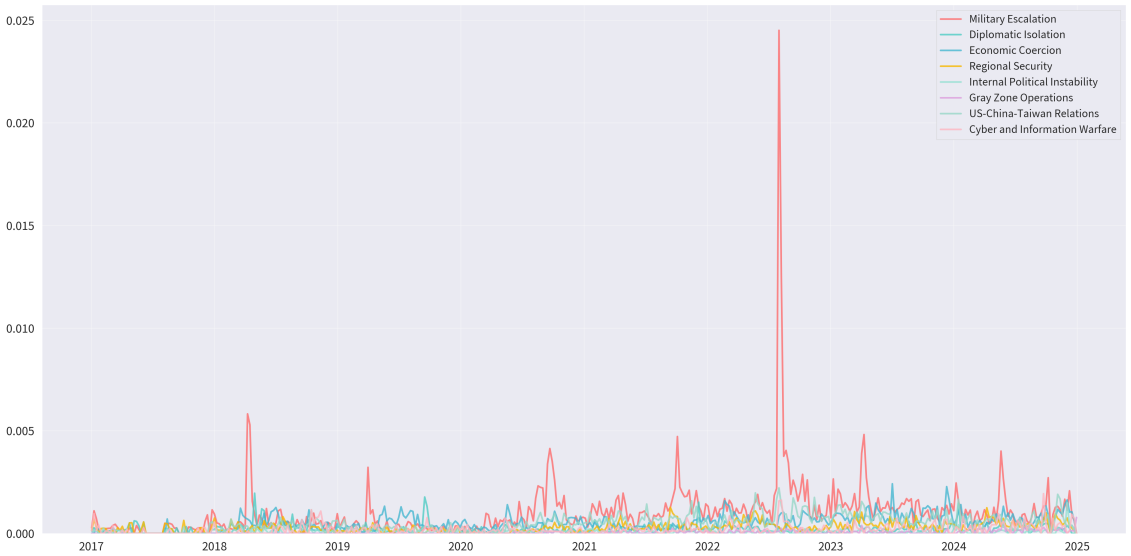


Figure 3.4: Taiwan GPR Index by Category

doi:doi:10.6342/NTU202503488

3.2 Unique-Spike Event Analysis

We then deep dive into unique spikes from our index and their underlying causes. Table 3.1 summarizes each spike. The summary shows that our index's unique strength lies in capturing multi-dimensional geopolitical risks beyond traditional military tensions. While both indices detect major military events like Pelosi's visit, our approach uniquely identifies: economic exclusion (IPEF framework), diplomatic isolation campaigns (Honduras, Nauru), technology warfare intersections (chip export controls), regional security realignments (Japan defense strategy, South Korea martial law), and strategic uncertainty from leadership transitions (Xi's third term, US elections).

Table 3.1: Major unique spikes in the GPR index

Peak week	Event	Description
2021-10-04	Cross-Strait military escalation	Record 38-39 daily PLA aircraft incursions into Taiwan's ADIZ marked unprecedented military pressure ahead of National Day celebrations; domestic debates questioned defense readiness while cognitive warfare concerns emerged alongside traditional military threats amid ongoing strategic competition.
2021-12-06	US-China strategic competition	Australian assessments of China's pandemic-weakened power coincided with Taiwan's contentious referendum campaigns and online influence operation controversies; Bob Dole's death recalled Trump-Tsai diplomatic breakthrough while defense strategists debated denial strategy amid Omicron variant emergence.
2022-05-23	IPEF exclusion, Taiwan	Biden's Indo-Pacific Economic Framework launch excluded Taiwan despite security commitments, exposing economic-security disconnect in American strategy; Australian Labor victory maintained firm China stance while Taiwan faced regional economic marginalization amid criticism of US inconsistency.
2022-10-10	Cross-Strait military escalation	Chinese military exercises intensified following US semiconductor export controls targeting China's technology sector; chip war dynamics intersected with military tensions while Russia-Ukraine conflict parallels reinforced Taiwan's vulnerability narrative and civil defense preparations.

Continued on next page

doi:doi:10.6342/NTU202503488

Table 3.1: Major unique spikes in the GPR index (Continued)

2022-10-24	Cross-Strait unification pressure	Xi Jinping's unprecedented third term confirmation signaled hardline Taiwan policy continuation amid technological decoupling and global recession fears; Ukraine war implications and US-China tech restrictions created multidimensional pressure on Taiwan's security and economic resilience.
2023-01-09	Cross-Strait military escalation	China's large-scale military drills responded to planned McCarthy-Tsai meeting while Japan's defense strategy explicitly mentioned Taiwan; supply chain restructuring accelerated as geopolitical blocs solidified around US-China competition amid broader deglobalization trends.
2023-03-13	Cross-Strait diplomatic isolation	Honduras diplomatic severance reduced Taiwan's allies to thirteen demonstrating Beijing's sustained isolation campaign; Russia-Ukraine war's second year heightened Taiwan invasion analogies while semiconductor diplomacy emerged as central survival strategy amid intensifying strategic competition.
2023-12-11	South China Sea tensions	Philippines-China confrontations in disputed waters escalated alongside Taiwan election interference concerns; US Congressional dysfunction over Ukraine aid exposed alliance reliability questions while technology competition and regional maritime tensions created complex security environment.
2024-01-15	Taiwan Election Fallout	Lai Ching-te's presidential victory triggered immediate Nauru diplomatic severance orchestrated by Beijing; electoral outcome demonstrated democratic resilience while China's swift retaliation through diplomatic isolation previewed intensified pressure campaign against incoming administration.
2024-11-04	US Presidential Election	Trump's potential return created strategic uncertainty for Taiwan policy amid deepening US partisan divisions; cross-strait relations faced unpredictability as transactional approaches threatened traditional security commitments while global flashpoints from Ukraine to Middle East compounded instability.
2024-12-02	South Korea political crisis	South Korea's martial law crisis exposed regional democratic fragility while US semiconductor export controls tightened on Chinese firms; Taiwan Strait military pressure persisted as authoritarian trends intersected with technology competition and ongoing Russia-Ukraine conflict implications.

3.3 Keyword & Category Decomposition

To double-check the validity of the index, we further analyze what category each spike captured, and what words contribute to those GPR index rises.

doi:doi:10.6342/NTU202503488

We first examine which words are most characteristic of spike periods. Table 3.2 shows the top 10 words that appear in at least half of the spike periods for each index, ranked by their frequency ratio. The frequency ratio is defined as the word's frequency during spike periods divided by its frequency across the entire dataset. Our GPR index captures more specific Taiwan-related geopolitical terms with higher frequency ratios compared to that of Lau's more general military terms.

As for each category's ratio to the overall index for each spike, which is shown in Figure 3.5, in general, no dominant category with some exceptions. For example, 2022 Pelosi's visit results in all-time high military escalation articles.

Table 3.2: Top words in spikes

Our GPR Index (21 spikes)			**Lau's GPR Index (22 spikes)**		
Word	Freq. Ratio	Spike Count	Word	Freq. Ratio	Spike Count
圍台軍演	7.25	11	封鎖	4.62	22
中國軍演	6.73	11	艦隊	2.99	22
台灣旅行法	6.14	11	共艦	2.72	16
外交打壓	5.85	14	海峽中線	2.52	20
第一擊	5.55	13	試射	2.51	22
政權更迭	5.52	15	演習	2.26	22
政治危機	4.93	17	開戰	2.12	22
外交休兵	4.91	12	侵台	2.02	22
主權尊嚴	4.77	11	軍艦	1.97	22
供應鏈風險	4.75	14	兩岸戰爭	1.96	12

doi:doi:10.6342/NTU202503488

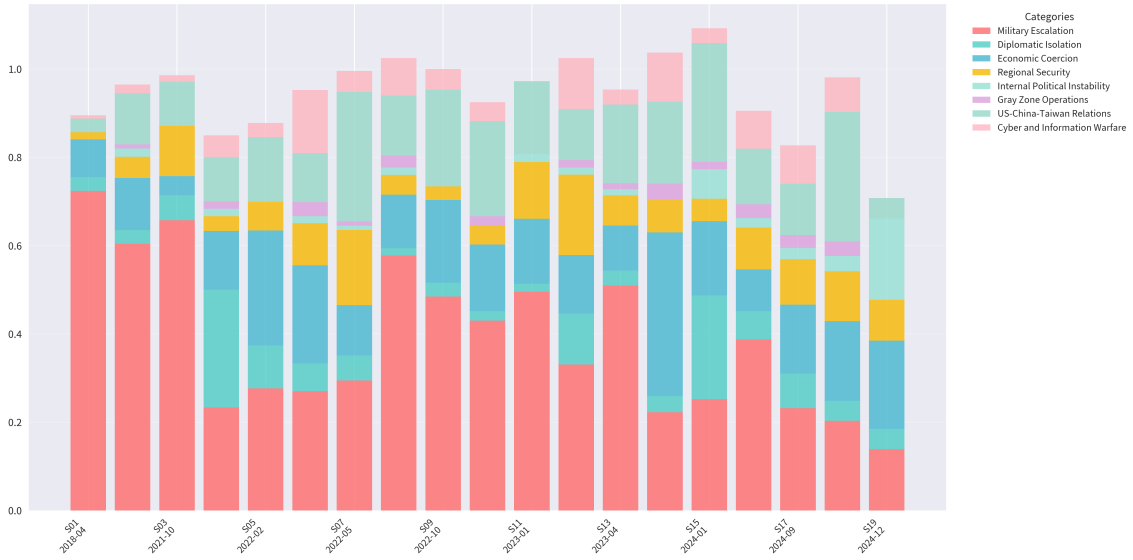


Figure 3.5: Category composition of GPR spikes

Each bar represents one spike event (S01–S19, sorted chronologically). The y-axis shows the ratio of category-specific GPR index to overall GPR index during that spike period.

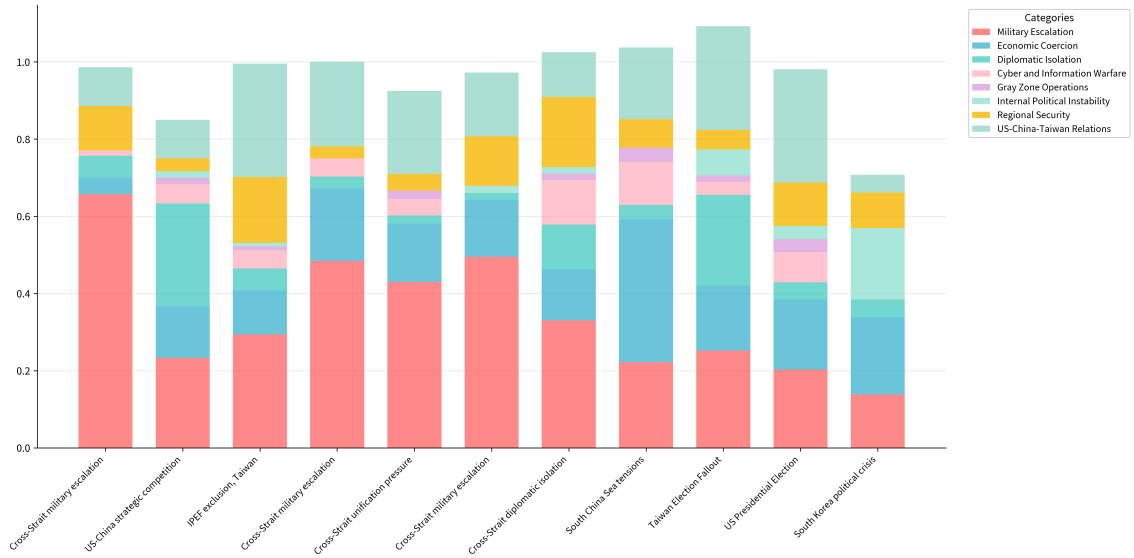


Figure 3.6: Proportion of each category for ch.5 unique spikes

doi:doi:10.6342/NTU202503488

3.4 Source-Heterogeneity Examination

This chapter analyzes source heterogeneity and its impact on Taiwan GPR index calculation. The GPR index by different media in Figure 3.7 shows that the behavior is quite diverse. Liberty Times tends to have a dramatic response to the GPR event and sometimes reaches zero related articles. However, China Times and United Daily have a consistent interest in GPR news while creating smaller spikes. We then look at all published article counts for each newspaper as shown in Figure 3.8. The large discrepancy indicates that each newspaper has a different content creation policy. The dramatic drop of Liberty Times news count may indicate the media's policy shift or the policy for bot crawling has changed mentioned in subsection 2.2.1. The consistently higher count of China Times articles actually comes from its policy to farm content, for example, daily summary of other newspapers.

Another evidence is the correlation score of GPR index constructed by different sources is much lower than Caldara and Iacoviello [2022]'s finding in US case. The Pearson correlation between China Times and Liberty Times is 0.475, between China Times and United Daily is 0.78 and between Liberty Times and United Daily is 0.559. They follow traditional view of their political position, from pro-KMT to pro-DPP, respectively are China Times - United Daily - Liberty Times. While Caldara and Iacoviello [2022] reported that between left and right newspaper, correlation is 0.87, Taiwan's media have larger position gap.[1]

To describe the gap, we list the top words that China Times and Liberty Times favor

[1] We also estimate it with enterprise traditional newspaper data on a small date range and using Lau's approach. The correlation between China Times and Liberty Times is 0.6, still significant lower than the US case.

doi:doi:10.6342/NTU202503488

differently in Table 3.3. We then choose some weeks that China Times and Liberty Times give large GPR category ratio differences, and summarize the differences in Table 3.4. It shows that Liberty Times frames China as an external threat and the U.S. as a partner. Liberty Times also sees drills as intimidation which will backfire on the CCP. On the other hand, China Times says the true danger is Taipei's independence line, and therefore drills are a predictable pushback.

We also construct a simple index which counts the ratio of articles mentioning ADIZ(防空識別區) and compares it with the counts of Taiwan's ADIZ violation by PLA in Figure 3.9. The figure shows that Taiwan's media did not follow the actual events. This is contrary to Caldara and Iacoviello [2022]'s finding that media's attention would not be attracted away in the long run trend. The figure also shows Liberty Times has a much higher interest in reporting ADIZ violation events which might come from its stance that views China as an external threat to justify Taiwan's independence.

The different media properties suggest that conducting Caldara and Iacoviello [2022]'s original method to treat all articles equally from different newspapers needs more consideration in Taiwan's case. The raw counts inconsistent make our GPR index depend more on China Times, resulting in a non-balanced aggregation. The high volatility of raw counts in short time periods makes the GPR index greatly affected by other fields' events. The fact that even one of the mainstream media in Taiwan also does content farming indicates that the denominator, count of all published articles, in the GPR index calculation needs better filtering in preprocessing, compared to Caldara and Iacoviello [2022]'s effort on molecular, GPR-related articles filtering.

doi:doi:10.6342/NTU202503488

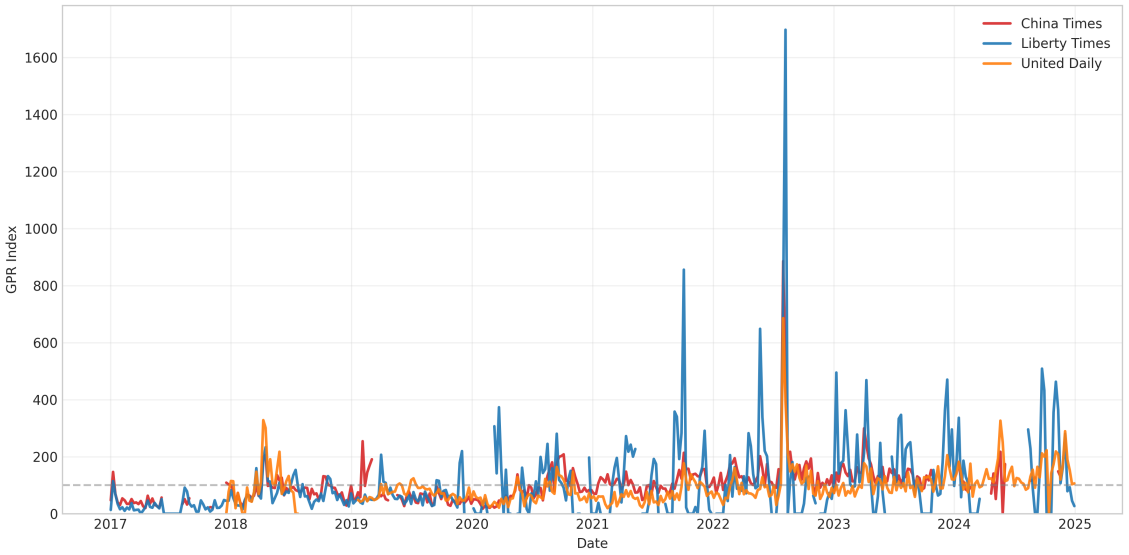


Figure 3.7: GPR Index for each newspaper

This figure shows Liberty Times has sharp spikes and sometimes no coverage, while China Times is more stable.

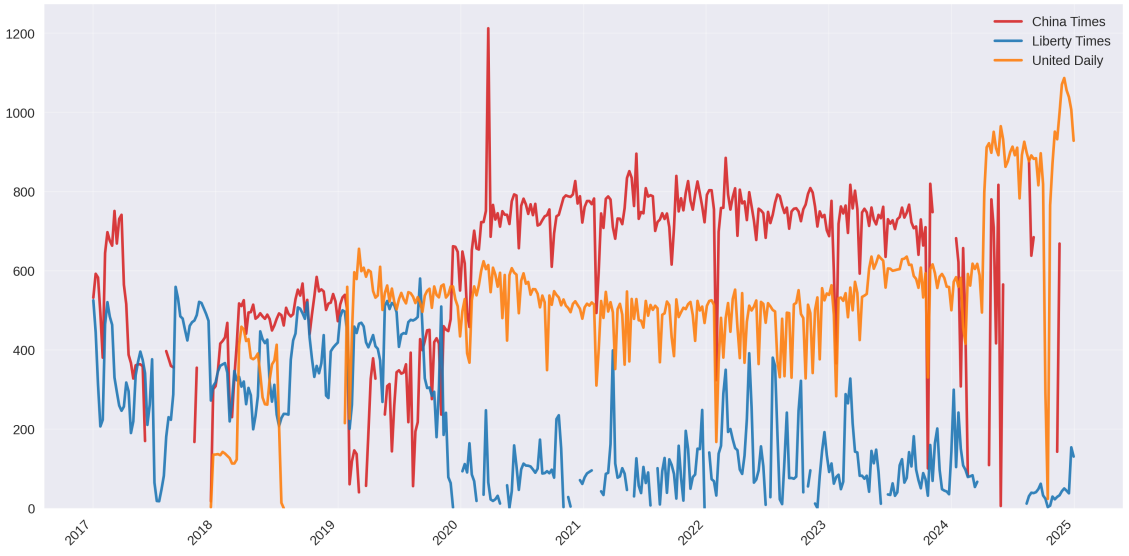


Figure 3.8: Articles count for each newspaper (daily average)

This figure confirms different publication volumes may bias the denominator in the index calculation.

doi:doi:10.6342/NTU202503488

Table 3.3: Relative word usage between China Times and Liberty Times

China Times (中國時報)		**Liberty Times (自由時報)**	
Word	Ratio†	Word	Ratio†
立法院衝突	0.27	台海局勢	1.34
金融制裁	0.56	召回大使	1.40
能源危機	0.62	日美安保	1.40
台海緊張	0.63	政局不穩	1.42
軍艦繞台	0.65	罷免	1.46
經濟制裁	0.67	外交斷交	1.46
台海危機	0.68	產業鏈轉移	1.48
群眾運動	0.68	網路滲透	1.49
飛彈試射	0.69	戰備警戒	1.53
軍演	0.69	輿論操控	1.70

Note. A ratio below 1 means the word appears proportionally less often in China Times than in Liberty Times; a ratio above 1 means the opposite.

†Ratio = *China Times mean□appearance rate ÷ Liberty Times mean□appearance rate*.

doi:doi:10.6342/NTU202503488

Table 3.4: Cross—Strait security framing in Liberty Times vs. China Times

Week	Liberty Times (pro-DPP)	China Times (pro-KMT)
2018-04-16	Beijing's live-fire drill framed as an external coercion; Premier Lai says "Taiwan is already independent," urging continued U.S. arms sales.	Drill portrayed as a justified warning to "pro-independence" politicians; editorials blame Lai for "lighting the fuse" and call for talks with Beijing.
2019-02-04	Election cycle; PLA video "only deepens Taiwanese disgust,"	warn DPP cannot escape independence problem
2022-03-14	Ukraine war used to stress U.S. support with stronger "porcupine" defences; urge even tougher conscription.	Ukraine cited to argue "relying on the U.S. is suicidal"; only abandoning independence keeps war away.
2023-08-21	PLA sea—air drill after Lai's U.S. transit cast as China's aggression; U.S. urges Beijing to stop pressure and resume dialogue.	Drills framed as Beijing's necessary response to "provocative" DPP moves; columns say PLA can "fight any time," urging Taipei to drop independence.

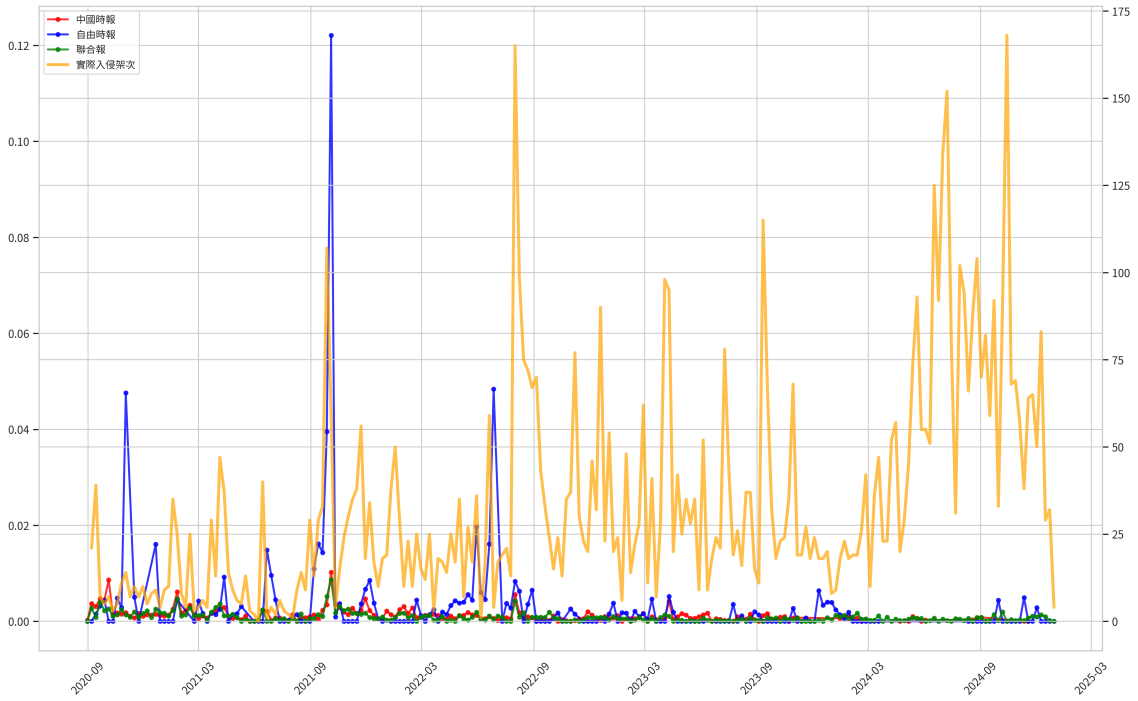


Figure 3.9: Articles ratio mentioned ADIZ vs counts of ADIZ invasion

doi:doi:10.6342/NTU202503488

3.5 VAR Analysis

In this section, we estimate the relationship between the GPR index and other macroeconomic indicators by using a structural VAR model. Follow Caldara and Iacoviello [2022] settings, the VAR model using two lags and quarterly data. And we slightly changed that making all variables stationary following Brignone et al. [2024]. Specifically, we take the log-difference transformation of real fixed private investment and real TWSE Capitalization Weighted Stock Index(TAIEX), take first differences of the policy rate, while party support rate and VIX remain in levels because they are already stationary. The log of private work hours per capita is seasonally adjusted considering Chinese New Year. The party support is the support rate of pro-DPP parties, since during our data range, DPP is the ruling party. Therefore, the support rate of pro-DPP parties represents the government's support rate.

The shock is obtained by using a Cholesky decomposition with the GPR index ordered first exactly as Caldara and Iacoviello [2022] did. Figure 3.10 shows the impulse response to a two standard deviation shock to the GPR index. The results are almost opposite to traditional expectations of GPR shock. Investment, work hours, TAIEX returns, and party support all respond to an increase; even VIX becomes lower in the short term. All of the above are opposite to Caldara and Iacoviello [2022]'s US findings. But it aligns with López Prol [2025]'s finding that Taiwan's market keeps strong while GPR increased. Another possible factor might be that monthly data is too vague to identify GPR effect as Baur and Smales [2020] found, aligning with Antonakakis et al. [2017]'s result.

The real cause needs further investigation; one way is to estimate a DCC model on daily GPR and market data, as Smales [2021] suggests.

doi:doi:10.6342/NTU202503488

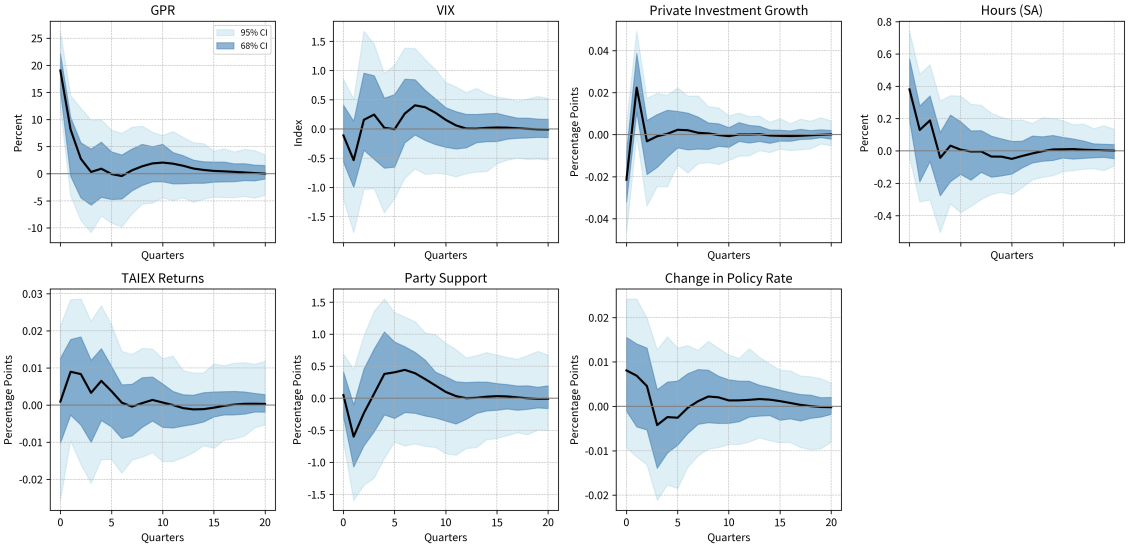


Figure 3.10: The Impact of Increased Geopolitical Risk

doi:doi:10.6342/NTU202503488

doi:doi:10.6342/NTU202503488

Chapter 4 Conclusion

This thesis asks whether large language models can reliably supply Taiwanese GPR dictionaries. By averaging the 40 indices implied by 40 LLM-generated dictionaries, we obtain a median series that sharpens the hybrid-warfare signal while sidestepping single-prompt idiosyncrasies. The ensemble cuts category imbalance, flags hybrid-tactic spikes that earlier lists missed, and correlates 0.61 with Lau [2024] despite methodological differences.

Decomposing the series by newspaper reveals substantial source heterogeneity; some outlets swing between silence and saturation, and web crawling gaps occasionally distort the article denominator. These findings raise concerns against treating press coverage as homogeneous, especially on online media data.

Linking the index to the macroeconomics yields a surprisingly resilient picture. A six-variable quarterly VAR indicates that a two standard deviation GPR shock is followed by an increase in investment, work hours, and equity returns, alongside a brief fall in implied volatility. Rather than contradicting the international literature, this pattern suggests that Taiwanese markets have partly internalised a chronic level of geopolitical tension and leaves open the question of how the stress affects micro-level firms, a possibility worth exploring with higher frequency data.

doi:doi:10.6342/NTU202503488

Beyond the Taiwan case, the exercise shows that very little domain expertise is required. The pipeline is essentially prompt, filter, count. Because the requirement of domain knowledge is outsourced to the model, the same workflow can target any topic, even one defined only by a fuzzy question. One could, for example, instruct the model to tag every article worldwide for "TSMC related geopolitical risk" and immediately obtain a cross-country series that tracks how threats to the foundry arise across time and space. The method also scales to a fully end-to-end classifier in which the LLM returns probabilities rather than an intermediate dictionary step, removing the dictionary step altogether.

Several limitations remain. Crawler coverage fluctuates with site policy changes; LLMs may be bad at earlier periods of historical data because nowadays data dominate their training data; and LLM outputs still are in need of careful manual review. Future work should experiment with daily financial linkages on a more accurate data source. Even so, the thesis demonstrates that with careful curation, large language models can assist traditional dictionary methods and offer a low-effort tool on the evolving research of geopolitical risk.

doi:doi:10.6342/NTU202503488

References

Nikolaos Antonakakis, Rangan Gupta, Christos Kollias, and Stephanos Papadamou. Geopolitical risks and the oil-stock nexus over 1899—2016. Finance Research Letters, 23:165–173, 2017. ISSN 1544-6123. doi: https://doi.org/10.1016/j.frl.2017.07.017. URL https://www.sciencedirect.com/science/article/pii/S1544612317300600.

Dirk G. Baur and Lee A. Smales. Hedging geopolitical risk with precious metals. Journal of Banking & Finance, 117:105823, 2020. ISSN 0378-4266. doi: https://doi.org/10.1016/j.jbankfin.2020.105823. URL https://www.sciencedirect.com/science/article/pii/S037842662030090X.

Davide Brignone, Luca Gambetti, and Martino Ricci. Geopolitical risk shocks: When the size matters. ECB Working Paper 2972, European Central Bank, August 2024. URL https://ssrn.com/abstract=4919668. Available at SSRN: https://ssrn.com/abstract=4919668.

Dario Caldara and Matteo Iacoviello. Measuring geopolitical risk. Technical Report 1222, International Finance Discussion Papers, 2018. URL https://doi.org/10.17016/IFDP.2018.1222.

doi:doi:10.6342/NTU202503488

Dario Caldara and Matteo Iacoviello. Measuring geopolitical risk. American Economic Review, 112(4):1194–1225, 2022. doi: 10.1257/aer.20191823.

Dario Caldara, Sarah Conlisk, Matteo Iacoviello, and Maddie Penn. Do geopolitical risks raise or lower inflation. Federal Reserve Board of Governors, pages 1–34, 2022.

Tzu-Chieh Hung and Tzu-Wei Hung. How china's cognitive warfare works: A frontline perspective of taiwan's anti-disinformation wars. Journal of Global Security Studies, 7 (4):ogac016, 07 2022. ISSN 2057-3170. doi: 10.1093/jogss/ogac016. URL https://doi.org/10.1093/jogss/ogac016.

Yeu-Tong Lau. 地緣政治風險如何影響台灣投資中國的上市公司股價. Master's thesis, 國立臺灣大學, Jan 2024.

Lincan Li, Jiaqi Li, Catherine Chen, Fred Gui, Hongjia Yang, Chenxiao Yu, Zhengguang Wang, Jianing Cai, Junlong Aaron Zhou, Bolin Shen, Alex Qian, Weixin Chen, Zhongkai Xue, Lichao Sun, Lifang He, Hanjie Chen, Kaize Ding, Zijian Du, Fangzhou Mu, Jiaxin Pei, Jieyu Zhao, Swabha Swayamdipta, Willie Neiswanger, Hua Wei, Xiyang Hu, Shixiang Zhu, Tianlong Chen, Yingzhou Lu, Yang Shi, Lianhui Qin, Tianfan Fu, Zhengzhong Tu, Yuzhe Yang, Jaemin Yoo, Jiaheng Zhang, Ryan Rossi, Liang Zhan, Liang Zhao, Emilio Ferrara, Yan Liu, Furong Huang, Xiangliang Zhang, Lawrence Rothenberg, Shuiwang Ji, Philip S. Yu, Yue Zhao, and Yushun Dong. Political-llm: Large language models in political science, December 2024. URL https://arxiv.org/abs/2412.06864.

Javier López Prol. Taiwan's stock market resilience to increasing geopolitical risk. Finance Research Letters, 80:107319, 2025. ISSN 1544-6123. doi: https://

doi:doi:10.6342/NTU202503488

doi.org/10.1016/j.frl.2025.107319. URL https://www.sciencedirect.com/science/article/pii/S1544612325005823.

Piotr Pietrzak. The xi jinping doctrine and the idea of china's peaceful coexistence in the fragile post-2022 context: A smokescreen or a long-term commitment? In Piotr Pietrzak, editor, Exploring the Implications of Local and Regional Conflicts, pages 257–290. IGI Global Scientific Publishing, 2025. doi: 10.4018/979-8-3693-2877-4.ch010. URL https://doi.org/10.4018/979-8-3693-2877-4.ch010.

Amna Sohail Rawat and Imtiaz Arif. Does geopolitical risk drive equity price returns of bric economies? evidence from quantile on quantile estimations. Journal of Finance & Economic Research, 3(2):24–36, June 2018. doi: 10.20547/jfer1803202. URL https://doi.org/10.20547/jfer1803202.

L.A. Smales. Geopolitical risk and volatility spillovers in oil and stock markets. The Quarterly Review of Economics and Finance, 80:358–366, 2021. ISSN 1062-9769. doi: https://doi.org/10.1016/j.qref.2021.03.008. URL https://www.sciencedirect.com/science/article/pii/S1062976921000533.

doi:doi:10.6342/NTU202503488

doi:doi:10.6342/NTU202503488

Appendix A — LLM Prompt Templates

This appendix presents the prompt templates used to generate GPR dictionaries through LLMs.

A.1 Step 1: Role-Based Dictionary Generation

The Step 1 prompt uses role-playing to generate diverse dictionaries. The system message varies the role (e.g., "Taiwan's defense minister", "foreign correspondent", "president of Taiwan") to obtain different perspectives on geopolitical risk.

Listing A.1: Step 1 Prompt Template

```
{
  "system": "You are {role}. Your task is to create a comprehensive
           dictionary of Traditional Chinese keywords relevant for
           identifying Geopolitical Risk (GPR) specifically related
           to Taiwan.",

  "user": "Please generate a Geopolitical Risk (GPR) dictionary
          focused exclusively on risks relevant to Taiwan, suitable
          for analyzing Taiwanese news media (using Traditional
          Chinese characters).

          Your response MUST be a single JSON object containing two
          top-level keys: `categories` and `word_sets`.

          1. **`word_sets`**: Define multiple lists of keywords under
             this key. Give each list a descriptive name related to
             a geopolitical risk theme relevant to Taiwan. The
             keywords within each list should be specific terms
             (nouns, verbs, proper nouns, common phrases) used in
             Taiwanese media to discuss that aspect of geopolitical
             risk.

          2. **`categories`**: Define GPR categories relevant to
             Taiwan under this key. Each category should be a
             key-value pair where the key is the category name and
             the value is a string defining the category logic by
             combining relevant `word_set` names using the `AND`
             operator (e.g., `theme_A AND theme_B`). You can combine
             any number of relevant word sets.
```


doi:doi:10.6342/NTU202503488

```
          Focus on terms and concepts frequently appearing in
          discussions of potential conflicts, threats, military
          buildups, diplomatic incidents, and other GPR events
          related to Taiwan."
}
```

A.2 Step 2: Fixed 8-Category Dictionary Generation

The Step 2 prompt enforces a stable 8-category structure across all dictionaries to ensure consistency while maintaining diversity in keyword selection.

Listing A.2: Step 2 Prompt Template

```
{
  "system": "You are an expert in Taiwan's geopolitical risks. Your
           task is to create a comprehensive dictionary of
           Traditional Chinese keywords for identifying Geopolitical
           Risk (GPR) specifically related to Taiwan.",

  "user": "Please generate a Geopolitical Risk (GPR) dictionary
          focused exclusively on risks relevant to Taiwan, suitable
          for analyzing Taiwanese news media (using Traditional
          Chinese characters).

          STRICT REQUIREMENTS:
          1. Create EXACTLY 8 categories - no more, no less
          2. Categories MUST be these 8 (you may use slight
             variations in naming but must cover these concepts):
             - Military Escalation (軍事升級) - military threats,
               exercises, deployments, cross-strait military
               tensions
             - Internal Political Instability (內部政治不穩定) -
               domestic unrest, political polarization, internal
               security threats
             - Economic Coercion (經濟脅迫) - sanctions, trade
               restrictions, supply chain pressure
             - Diplomatic Isolation (外交孤立) - international
               recognition challenges, diplomatic pressure
             - Cyber and Information Warfare (網路與資訊戰) - cyber
               attacks, disinformation, cognitive warfare
             - US-China-Taiwan Relations (美中台關係) - trilateral
               dynamics, strategic competition
             - Regional Security (區域安全) - Asia-Pacific tensions,
               alliance dynamics
             - Gray Zone Operations (灰色地帶行動) - hybrid warfare
               below war threshold

          3. **`word_sets`**: Define multiple lists of keywords under
             this key. Give each list a descriptive name related to
             the risk themes. Each list should contain 15-25
             Traditional Chinese terms/phrases used in Taiwanese
             media.

          4. **`categories`**: Each category should combine relevant
             word_set names using ONLY the `AND` operator (e.g.,
             `military_threats AND cross_strait_tensions`). Do NOT
             use OR operators.

          5. Ensure diversity in term selection across runs while
             maintaining relevance to Taiwan's context."
}
```


doi:doi:10.6342/NTU202503488

Both prompts use JSON schema validation to ensure structured output with the required ‘categories‘ and ‘word_sets‘ keys. The Step 2 prompt additionally enforces exactly 8 categories and 15-25 words per word set through schema constraints.

doi:doi:10.6342/NTU202503488

doi:doi:10.6342/NTU202503488

Appendix B — Supplementary Tables / Code Links

This appendix provides supplementary tables and links to the code.

B.1 code link

https://github.com/daniel-bale/gpr-zero

B.2 Category Mapping Analysis: Consolidating GPR Categories into 8 Representative Groups

This section analyzes how the 172 unique categories found across all GPR dictionaries can be mapped to 8 representative categories. This consolidation enables stable prompt generation while maintaining comprehensive coverage of geopolitical risk dimensions.

doi:doi:10.6342/NTU202503488

B.2.1 Rationale for Category Selection

B.2.1.1 Why "Cross-Strait Political Tensions" is NOT a separate category

In recent years, cross-strait political and diplomatic tensions rarely occur in isolation - they almost always come with military escalation. Pure political rhetoric without military backing does not constitute a significant geopolitical risk in the current environment. Additionally, political tensions alone don't qualify as hybrid warfare and are difficult to justify as a standalone risk category. Therefore, cross-strait political elements are incorporated into Military Escalation where they manifest as actual threats.

B.2.2 The 8 Representative Categories

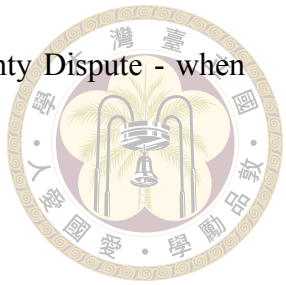
B.2.2.1 1. Military Escalation

Direct military threats, exercises, deployments, potential for armed conflict, AND cross-strait military tensions.

Mapped Categories: Military Escalation, MilitaryConflictRisk, Military_Escalation, Military_Tensions, Direct_Military_Confrontation, Military Conflict, 軍事威脅 (Military Threat), 軍演威嚇 (Military Exercise Intimidation), 軍演恫嚇 (Military Exercise Coercion), direct_military_action, direct_military_confrontation_risk, direct_military_threat, military_conflict, military_exercise_or_threat, military_risk, military_tensions, threat_of_invasion, 直接軍事衝突風險 (Direct Military Conflict Risk), CrossStraitTensions (when military-related), Cross_Strait_Political_Deterioration (when leads to military action), cross_strait_conflict, cross_strait_military_escalation, cross_strait_political_standoff (when involves military posturing), cross_strait_political_tension (when involves military threats), 台海危機 (Tai-

doi:doi:10.6342/NTU202503488

wan Strait Crisis), 海峽危機 (Strait Crisis), 主權爭議 (Sovereignty Dispute - when contested militarily), sovereignty_dispute_escalation.



B.2.2.2 2. Internal Political Instability

Domestic political risks, internal unrest, and security threats within Taiwan.

Mapped Categories: Internal Instability, Internal Political Instability, Internal_Political_Instability, Internal_Political_Stability, domestic_instability, domestic_political_risk, domestic_political_risks, domestic_unrest, 內部安全威脅 (Internal Security Threat), 內部滲透 (Internal Infiltration), internal_sovereignty_debate_under_pressure, TaiwaneseElectionsRisk, Humanitarian_Crises.

B.2.2.3 3. Economic Coercion

Economic pressure, sanctions, trade restrictions, and supply chain vulnerabilities.

Mapped Categories: Economic Coercion, Economic Pressure, EconomicSanctionsRisk, Economic_Coercion, Economic_Coercion_and_Supply_Chain_Risk, Economic_Pressure, economic_and_technological_coercion, economic_coercion, economic_coercion_event, economic_coercion_incident, economic_risk, economic_sanctions_event, economic_security_incident, 經濟脅迫 (Economic Coercion), 經濟與科技安全威脅 (Economic and Tech Security Threats), 國際制裁 (International Sanctions), 國際制裁風險 (International Sanctions Risk), tech_supply_chain_risk, geopolitical_supply_chain_disruption_risk, supply_chain_disruption_risk.

B.2.2.4 4. Diplomatic Isolation

International recognition challenges, diplomatic pressure, and isolation campaigns.

doi:doi:10.6342/NTU202503488

Mapped Categories: Diplomatic Crisis, Diplomatic Crisis manual, Diplomatic Isolation, Diplomatic_Conflicts, Diplomatic_Isolation, Diplomatic_Isolation_Campaign, diplomatic_coercion_event, diplomatic_conflicts, diplomatic_contention, diplomatic_isolation, diplomatic_isolation_crisis, diplomatic_isolation_incident, diplomatic_isolation_pressure, diplomatic_pressure, diplomatic_pressure_and_isolation, diplomatic_pressure_incident, international_diplomatic_pressure, 外交孤立 (Diplomatic Isolation).

B.2.2.5 5. Cyber and Information Warfare

Digital threats, cyber attacks, disinformation campaigns, and cognitive warfare.

Mapped Categories: Cyber & Disinformation Threats, Cyber & Information Warfare, Cyber and Disinformation Threat, Cyber and Disinformation Threats, Cyber and Information Warfare, Cyber/Information Warfare, CyberSecurityThreats, Cyber_Threats, Cybersecurity_Threats, cybersecurity_risk, cybersecurity_risks, cybersecurity_threats, information_warfare, cognitive_and_cyber_warfare, 網路攻擊 (Cyber Attack), 網路與資訊戰 (Cyber and Information Warfare), 認知作戰 (Cognitive Warfare), internal_disinformation_campaign.

B.2.2.6 6. US-China-Taiwan Relations

Trilateral dynamics, strategic competition, and great power involvement.

Mapped Categories: US-China-Taiwan Strategic Risk, US-China-Taiwan Tensions, USChinaTaiwanRelations, US_China_Taiwan_Dynamics, US_China_Tension_over_Taiwan, us_china_confrontation_over_taiwan, us_china_escalation_over_taiwan, us_china_standoff_over_taiwan, us_china_taiwan_tensions, us_taiwan_china_triangular_friction, Major Powers Involvement, Foreign Intervention Risk, foreign_defense_support, foreign_military_involvement,

doi:doi:10.6342/NTU202503488

US-Taiwan Security Dynamics, US_Taiwan_Support, us_defense_support_and_arms_sales, us_diplomatic_interaction, us_intervention_signal, us_security_support_for_taiwan, us_taiwan_security_co

B.2.2.7 7. Regional Security

Broader Asia-Pacific tensions, regional conflicts, and alliance dynamics.

Mapped Categories: Regional Security Tensions, Regional_Security_Risk, 區域緊張 (Regional Tensions), 區域衝突 (Regional Conflict), 區域軍事化 (Regional Militarization), 區域軍備競賽 (Regional Arms Race), regional_alliance_activity, regional_alliance_response_to_threat, regional_conflicts, regional_escalation, regional_instability, regional_security_destabilization, 安全聯盟 (Security Alliance), 聯盟動態 (Alliance Dynamics), 聯盟關係 (Alliance Relations), InternationalSupportForTaiwan, International_Alliance_Response_to_Threat.

B.2.2.8 8. Gray Zone Operations

Hybrid warfare, coercive activities below war threshold, and ambiguous threats.

Mapped Categories: Gray Zone Operations, Gray Zone Pressure, Gray_Zone_Conflict, gray_zone_activity, gray_zone_escalation, gray_zone_warfare, grey_zone_conflict, grey_zone_conflict_esc 灰色地帶行動升級 (Gray Zone Operations Escalation), 灰色地帶衝突 (Gray Zone Conflict), Hybrid_Warfare_Campaign.

B.2.3 Additional Category Mappings

Infrastructure/Energy → Economic Coercion or Military Escalation: Critical Infrastructure Risk, Critical Infrastructure Threat, Critical Infrastructure Threats, Criti-

doi:doi:10.6342/NTU202503488

cal_Infrastructure_Risk, Energy_Security_Risks, energy_security, 能源安全 (Energy Security), 封鎖威脅 (Blockade Threat).

Technology/Science → Economic Coercion or Cyber Warfare: 科技戰爭 (Technology War).

Defense Readiness → Military Escalation or Regional Security: taiwan_defense_posturing, taiwan_defense_preparedness, taiwan_defense_readiness, taiwan_defense_readiness_change, 台灣國防整備與改革 (Taiwan Defense Preparation and Reform).

B.2.4 Validation Summary

The 8 representative categories successfully capture all major dimensions of geopolitical risk:

- **Coverage**: All 172 unique categories can be mapped to one or more of the 8 categories
- **Balance**: Each representative category has substantial mapped subcategories
- **Clarity**: The categories are distinct enough to avoid excessive overlap
- **Comprehensiveness**: Together they cover military, political, economic, diplomatic, technological, and strategic dimensions

doi:doi:10.6342/NTU202503488