

Big Data

The BBVA Research Geopolitics Monitor: Tracking Geopolitical Sentiment and Events using Natural Language Techniques

Alvaro Ortiz (BBVA Research) / Tomasa Rodrigo (BBVA Research) / Jorge Sicilia (BBVA Research)

November 29, 2023

Abstract

During the last decade we have witnessed a multidimensional digital transformation characterized by the increasing power of computers and a fast development of new algorithms, accompanied by a vast amount of data. The capacity of new algorithms to convert these mostly unstructured data (text, images, video, audio, satellites) into data ready to be analyzed is changing the way in which we can approach many social sciences. In particular, the recent development of Large Language Models has the capacity to change the way we do research, gaining scale and scope. In this brief we show how at BBVA Research we convert text from news to obtain different numeric indicators that quantify what and how the media feels about the geopolitical events using Natural Language Processing (NLP) techniques that are improving our capacity to understand the world.

Keywords: Machine Learning, Unstructured Information, Geopolitics, Conflicts, Big Data.

Big Data and Geopolitics

The increasing power of computers, the development of algorithms - particularly those related to Natural Language - and the vast volume of unstructured data (text, video, audio, satellite...) brought about by digitalization has triggered a rapid advance of the Artificial Intelligence applications in social sciences, including economics, politics and international relations; thus also geopolitics. In the field of international relations, the advantage is significant as the quantitative information was scarcer.

Extracting unstructured information from “news” to monitor geopolitical events is not new. The first tests began around 2010 with the development of the Integrated Crisis Early Warning System (ICEWS)¹. Part of the model includes a database of political events coded by experts and included as a part of a system to provide conflict early warnings.

At present, the use of NLP techniques allows us to track multiple events, providing answers to increasingly complex questions revolving around a wide range of emotions. At BBVA Research we have been using these methods for almost a decade as we saw with some frustration the shortcomings of only using storytelling to understand conflicts that could transform into risks affecting the economic outlook. This technique has provided

1: ICEWS has recently morphed to POLECAT, a more advanced version. ICEWS uses a mixed methods approach to instability forecasting, combining heterogeneous statistical and agent-based models in an integration framework with an aggregate forecast accuracy of more than 80%. These models are provisioned in near real-time from more than 100 data sources and 250 international and regional news feeds. Millions of news stories are processed by an innovative, shallow-parsing Jabari technology and BBN's Serif NLP technology.

some answers almost in real time and in a very granular way. A long way has already passed since we started using it to follow the Arab Spring.

The combination of Big Data and geopolitics is a burgeoning field, offering vital tools to help navigate the complex and interconnected world of international relations.

“Text to Numbers” using Natural Language to analyze Geopolitics

In BBVA Research we have developed a broad range of Big Data sentiment indicators to track the geopolitical and social spheres using Big Data coming from news through NLP techniques. The raw source where all the data comes from is the [Global Database of Events, Language and Tone \(GDELT\)](#)², which is an open-source database that extracts and parses digital news in broadcast, print and web media globally in over 100 languages on a daily basis, from global to local media sources.

GDELT uses different dictionaries, and several thousands of taxonomies are identified in the news pieces to classify and categorize the information. The algorithms can identify organizations, locations, news sources and events across the world as well as emotions and sentiments³. Every processed socio-political event in GDELT is coded using the [Conflict and Mediation Event Observations \(CAMEO\)](#) event coding system developed by Gerner et al. (2002)⁴.

To build all the BBVA Research “social” indexes we capture both the coverage and sentiment of news articles per day including any mention of whatever topic we want to monitor. Thus, we have developed a large set of indexes on a broad set of areas including geopolitical risks⁵, political stability⁶, conflict and protest indexes, and indexes on economic policy uncertainty⁷ and bilateral sentiment between countries. Many of them are potentially temporarily linked: a protest in a country can turn into a conflict or into higher economic policy uncertainty; some move across countries: protests spread across different areas. Some indexes show different readings when filtering for foreign or local media sources and can thus be used to explain different sets of variables.⁸

All our indexes are normalized so they show their relative performance to its own past, what we define “signals”, and we apply them to a geometric moving average of 28 days to reduce noise but give more weight to the more recent daily information. The next section explains with a bit more detail a particular example of a few of our geopolitical indexes.

2: For detailed information about GDELT, you can consult www.gdelt.org.

3: Focusing on sentiment, once each news piece is translated into English, GDELT applies more than 40 different dictionaries that classify words associated with positive and negative tone to compute the average “tone” of all documents containing one or more mentions to the events we are looking for. The score ranges from -100 (extremely negative) to +100 (extremely positive).

4: CAMEO is a broadly used coding scheme to systematize the analysis of political and social events and divide them in a scale ranging from material and verbal cooperation to verbal and material conflict.

5: In the case of the BBVA Geopolitical Risk index, we follow the methodology of Caldara and Iacovello (2022) to identify the intersection of two sets of keywords related to geopolitical risk.

6: For political tensions, we collect all news articles that mention politics related topics following the GDELT taxonomy such as elections and campaigns, political parties and politicians, government institutions, policies, political scandals etc.

7: The economic policy uncertainty (EPU indicator is based on the Baker, Bloom & Davis (2016) using tone and relative coverage associated with keywords related to Economic Uncertainty and monetary, fiscal and regulatory policy.

8: Economic policy uncertainty indicators from local media typically anticipate better economic perspectives than those from external media; but the latter are better able to anticipate how the exchange rate evolves.

How the World Feels about the Israel-Hamas Conflict and the recent ease of US and China relations?

The world has been recently exposed to several geopolitical events which have stressed the World Order. In this section we show how some of our sentiment indexes have captured two different sets of events: the impact of the recent Israel-Hamas Conflict in individual countries and the relationship between China and the US, using bilateral links.

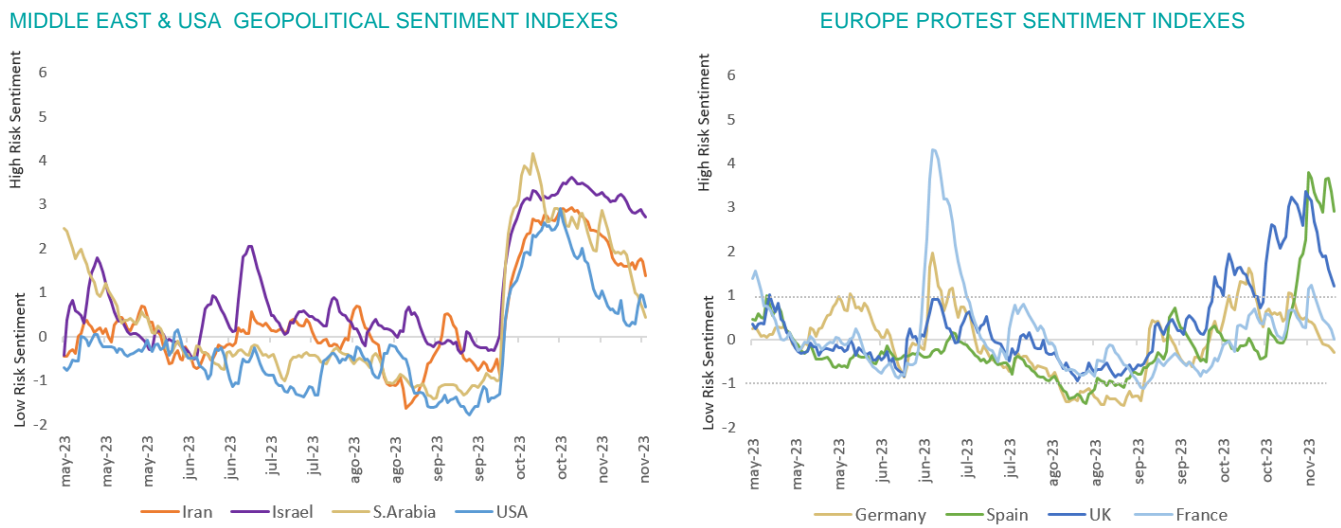
The Israel-Hamas Conflict

When Hamas attacked Israel on October 7th the world geopolitical media sentiment was somehow calmer relative to the recent past. While the Russia-Ukraine conflict was well alive, as shown on our map of signals at the beginning of October, most of the world seemed to be set in a relatively calm situation.

Graph 1 displays the geopolitical sentiment of the Middle East countries and the US. As shown in the first graph, all indexes jumped immediately to reach extreme risk levels and remained at those high levels in all countries of the sample except the US, whose index is clearly curbing towards a more neutral sentiment as news started to digest the initial surprise and other pieces of news have probably taken the center stage.

The spill-overs effects from the conflict to Europe were also evident since the beginning of the crisis with protests erupting in many European countries. Sentiment polarized and people began to protest in the streets. This has been particularly relevant in the UK and, to a lesser extent, in France. In other countries the impact was lower⁹.

Figure 1. **THE GEOPOLITICAL SENTIMENT IN THE MIDDLE EAST & PROTEST SENTIMENT IN WESTERN COUNTRIES** (SIGNALS SCALED ACCORDING TO STANDARD DEVIATIONS OF INDIVIDUAL SERIES)



Source: BBVA Research and GDELT

9: It jumped significantly in Spain but for reasons linked to domestic political protest (not shown in the graph) unrelated to the Israel Hamas conflict.

The following diagrams in Graph 2 show how the different sentiment indexes including Geopolitical Risk, Political tensions, Economic Policy Uncertainty and Protest and Conflict evolved since the beginning of the conflict:

- **The geopolitical risk, political instability, conflict, and protest increased in the Middle East to high or extreme levels.** One important exception has been the protest sentiment in Iran, which remained neutral. This is particularly important as some analysts were naturally concerned about the possibility of the crisis to mobilize Iranian people and destabilize the regime.
- **The crisis spilled over rapidly to some indexes in Western countries:** the Geopolitical and Conflict indicators moved rapidly to the high-risk area in some European Countries, particularly in the UK and Spain, where protests increased for local reasons.
- **Fortunately, the economic policy uncertainty index has remained resilient for now. This has been a broad-based pattern.** There are some explanations here. First, the Conflict language has displaced the economic one. Second, the Conflict has coincided in time with the key Central Banks maintained on the sidelines after a very aggressive tightening cycle and still resilient economic conditions, which has also spurred a very positive momentum in the stock market. Finally, the reaction of commodities has been muted, together with a lack of discussion of policy measures or major disruptions in any country outside those in conflict.

Figure 2. **THE BBVA RESEARCH GEOPOLITICS SENTIMENT MONITOR DURING THE HAMAS-ISRAEL CRISIS** (SIGNALS IN GEOPOLITICAL RISK, POLITICAL TENSIONS, ECONOMIC POLICY UNCERTAINTY, PROTEST AND CONFLICT. SIGNALS SCALED ACCORDING TO STANDARD DEVIATIONS OF INDIVIDUAL SERIES)

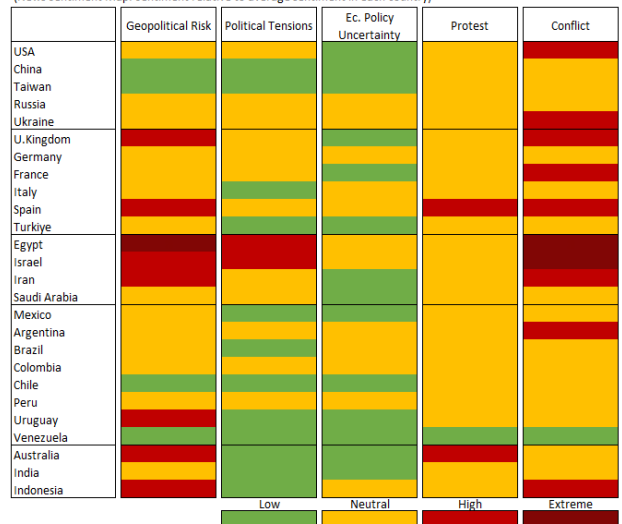
Summary of Signals Before Hamas-Israel Conflict (October 5th)

(News Sentiment Map. Sentiment relative to average sentiment in each country)



Summary of Signals After Hamas-Israel Conflict (November 29th)

(News Sentiment Map. Sentiment relative to average sentiment in each country)



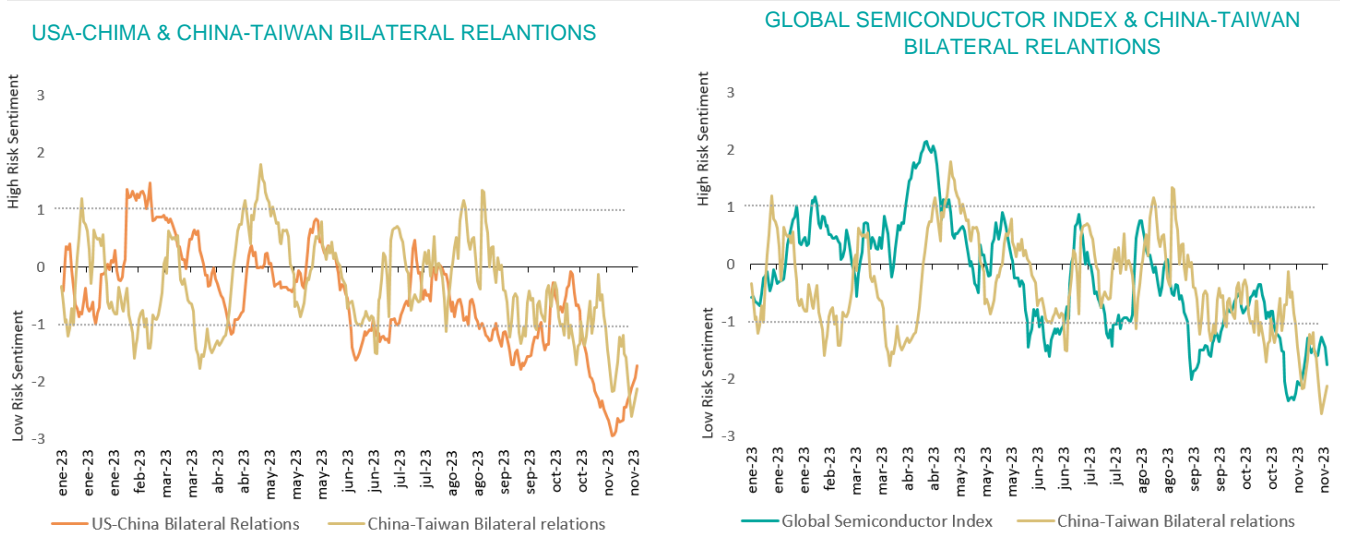
Source: BBVA Research and GDELT

The Bilateral Relations Index: China - US relations and the role of Taiwan

The bilateral relations index serves as an analytical tool to gauge the dynamics between two countries, detailing actions ranging from cooperation to conflict and pinpointing the actors involved. The index tracking US-China relations reflects a recent de-escalation of geopolitical tension, moving from high risk at the beginning of the year to the lowest risk level in two years following the preparatory talks for the Biden-Xi meeting at APEC. This thaw in relations mirrors a similar trend in the pressure exerted by China on Taiwan, which has maintained a more neutral stance but showed signs of easing following the US-China leadership meeting.

Following this index, one could conclude that for some time now the tensions have eased, following a series of policy meetings between high level political figures in both countries, which have been pointing for some time that both areas were trying to bring down tensions. This was probably indicating that the potential economic impact would be lower than what could be expected 8-10 months ago, particularly in sectors tied to the geopolitical climate and in the line of fire between the two countries. A notable example is the semiconductor industry, whose sentiment index shows a strong correlation with the China-Taiwan relationship. This makes a lot of sense considering Taiwan's central role in the semiconductor supply chain through TSMC (Taiwan Semiconductor Manufacturing Company)¹⁰.

Figure 3. **THE BBVA RESEARCH GEOPOLITICS SENTIMENT MONITOR DURING THE HAMAS-ISRAEL CRISIS** (SIGNALS IN GEOPOLITICAL RISK, POLITICAL TENSIONS, ECONOMIC POLICY UNCERTAINTY, PROTEST AND CONFLICT. SIGNALS SCALED ACCORDING TO STANDARD DEVIATIONS OF INDIVIDUAL SERIES)



Source: BBVA Research and GDELT

10: Forthcoming in Ortiz, A., Rodrigo, Tsu J and Villarta, P (2023): "A Sentiment analysis of the Semiconductor Crisis". Mimeo.

Conclusion

The BBVA Research Geopolitics Monitor shows how the application of advanced computational techniques can be used in the realms of economics and geopolitics. By harnessing the power of Natural Language Processing (NLP) and Large Language Models, this tool effectively translates the vast, unstructured data from global news sources into quantifiable, interpretable insights. The methodology, whose usefulness can be shown through case studies like the Israel-Hamas conflict and US-China relations, provides a nuanced understanding of geopolitical sentiments and events, crucial for policymakers, economists, and stakeholders in the global arena. Working with a large set of indexes in different countries and bilateral relations allows an analyst to quickly identify relevant events and the importance of geopolitical issues.

In essence, the Monitor acts as a barometer, capturing the ebbs and flows of geopolitical sentiments in real-time. This innovative approach marks a pivotal shift in how geopolitical risks and relationships are analyzed, allowing for a more proactive and informed stance in policy-making and economic strategy. The integration of Big Data and NLP in geopolitics not only enhances our understanding of global events but also paves the way for more sophisticated, data-driven decision-making processes in the interconnected world of today. Introducing these indexes in standard econometric frameworks allows us to simulate impact of different events on the evolution of economic and financial variables.

References

Scott R. Baker, Nicholas Bloom, Steven J. Davis, Measuring Economic Policy Uncertainty, *The Quarterly Journal of Economics*, Volume 131, Issue 4, November 2016, Pages 1593–1636,

Caldara, Dario, and Matteo Iacoviello. 2022. "Measuring Geopolitical Risk." *American Economic Review*, 112 (4): 1194-1225.

Gerner, Deborah & Jabr, Rajaa & Schrod, Philip. (2002). Conflict and Mediation Event Observations (CAMEO): A New Event Data Framework for the Analysis of Foreign Policy Interactions.

Halterman, A, Bagozzi, B., Beger, A., Schrod, P. Scarborough, G (2023). PLOVER and POLECAT: A New Political Event Ontology and Dataset. *Arxiv Social Science*

Ortiz, A., Rodrigo, Tsu J and Villarta, P (2023): "A Sentiment analysis of the Semiconductor Crisis". Mimeo.

DISCLAIMER

The present document does not constitute an “Investment Recommendation”, as defined in Regulation (EU) No 596/2014 of the European Parliament and of the Council of 16 April 2014 on market abuse (“MAR”). In particular, this document does not constitute “Investment Research” nor “Marketing Material”, for the purposes of article 36 of the Regulation (EU) 2017/565 of 25 April 2016 supplementing Directive 2014/65/EU of the European Parliament and of the Council as regards organisational requirements and operating conditions for investment firms and defined terms for the purposes of that Directive (MIFID II).

Readers should be aware that under no circumstances should they base their investment decisions on the information contained in this document. Those persons or entities offering investment products to these potential investors are legally required to provide the information needed for them to take an appropriate investment decision.

This document has been prepared by BBVA Research Department. It is provided for information purposes only and expresses data or opinions regarding the date of issue of the report, prepared by BBVA or obtained from or based on sources we consider to be reliable, and have not been independently verified by BBVA. Therefore, BBVA offers no warranty, either express or implicit, regarding its accuracy, integrity or correctness.

This document and its contents are subject to changes without prior notice depending on variables such as the economic context or market fluctuations. BBVA is not responsible for updating these contents or for giving notice of such changes.

BBVA accepts no liability for any loss, direct or indirect, that may result from the use of this document or its contents.

This document and its contents do not constitute an offer, invitation or solicitation to purchase, divest or enter into any interest in financial assets or instruments. Neither shall this document nor its contents form the basis of any contract, commitment or decision of any kind.

The content of this document is protected by intellectual property laws. Reproduction, transformation, distribution, public communication, making available, extraction, reuse, forwarding or use of any nature by any means or process is prohibited, except in cases where it is legally permitted or expressly authorised by BBVA on its website www.bbvaresearch.com.

ENQUIRIES TO:

BBVA Research: Azul Street, 4. La Vela Building – 4th and 5th floor. 28050 Madrid (Spain).
Tel. +34 91 374 60 00 y +34 91 537 70 00 / Fax (+34) 91 374 25
www.bbvaresearch.com