

MASTER'S DEGREE
ECONOMICS OF BUSINESS AND STRATEGY

The Influence of Media on Firms' Perception of Geopolitical Risk: The Portuguese Case

Pedro Daniel de Deus Carvalho

M

2026



THE INFLUENCE OF MEDIA ON FIRMS' PERCEPTION OF
GEOPOLITICAL RISK: THE PORTUGUESE CASE

Pedro Daniel de Deus Carvalho

Dissertation
Master in Economics of Business and Strategy

Supervised by
Professor José Manuel Peres Jorge

2026

Acknowledgments

I would like to express my sincere gratitude to the Observatório do Risco Geopolítico para Empresas for trusting me with the results and details of their Barometer, which were essential to the development of this thesis.

I am also deeply grateful to Professor José Magano, Professor Domingos Rodrigues and Burcu Degerli for their availability, patience and valuable guidance in helping shape this work.

To my family and friends, thank you for your never-ending support, encouragement and understanding throughout this journey.

Finally, I would like to extend my heartfelt thanks to my supervisor, Professor José Jorge, for all his help, insights and guidance, which were invaluable in bringing this thesis to completion.

Abstract

This dissertation analyzes the relationship between Portuguese media coverage of Geopolitical Risk and Portuguese firms' perception of that risk. Specifically, it aims to understand whether the Geopolitical Risks that received more coverage (Volume) or stronger affective tone (Sentiment) by Portuguese media were also the Geopolitical Risks that Portuguese companies perceived as the most relevant. Special attention will also be given to the way some firms characteristics (such as size or level of openness to foreign markets) and time horizon (short [1 year] and long term [3 years]) affect that relationship.

To achieve this, the results from the **Barómetro de Risco Geopolítico para Empresas 2025** (that measured the firms' perception of Geopolitical Risk) were combined with the results from the **Media-Weighted Geopolitical Risk Index** (that measured the media coverage of Geopolitical Risk and which was specifically constructed for this study). The results suggest that Volume is not aligned with corporate perceptions of Geopolitical Risk, while Sentiment shows stronger alignment, especially in larger and more internationally open firms.

Although previous contributions highlighted the importance of Geopolitical Risk for companies, existing literature failed to understand what factors influence the firms' perceptions. By demonstrating that media attention is one of the factors that explain corporate Geopolitical Risk perception, this dissertation contributes to a better understanding of risk perception, Geopolitical Risk and media influence that can be applied for Portuguese companies and foreign markets alike. Additionally, the methodology followed is, in itself, a relevant contribution to the literature, due to the use of LLM to interpret text.

Keywords: Geopolitical Risk; Geopolitical Risk Perception; Corporate Risk Perception; Media Coverage; Media Framing; Portuguese Firms; BRGE2025

Resumo

Esta dissertação ambiciona analisar a relação entre a cobertura a Risco Geopolítico feita pelos media portugueses e a perceção das empresas portuguesas desse risco. Mais especificamente, pretende perceber se os Riscos Geopolíticos que receberam maior cobertura (Volume) ou tom afetivo mais intenso (Sentiment) correspondem aos Riscos Geopolíticos percebidos como mais sérios pelas empresas portuguesas. Será também dada especial atenção à forma como determinadas características de empresas (como o tamanho ou o grau de abertura a mercados internacionais) e variações de horizonte temporal (curto [1 ano] e longo prazo [3 anos]) afetam essa relação.

De modo a atingir este objetivo, os resultados do Barómetro do Observatório do Risco Geopolítico para Empresas 2025 (que mediu a forma como as empresas portuguesas percecionaram o Risco Geopolítico) foram cruzados com os resultados do Media-Weighted Geopolitical Risk Index (que mediu a cobertura de Riscos Geopolíticos feita pelos media portugueses e que foi especificamente construído para esta dissertação). Os resultados sugerem que o Volume não está relacionado com as perceções das empresas, enquanto o Sentimento apresenta uma correlação mais forte, particularmente para grandes empresas e para empresas mais abertas aos mercados internacionais.

Apesar de prévias contribuições destacarem a relevância do Risco Geopolítico para as empresas, a literatura existente ainda não compreende que fatores influenciam a perceção de Risco Geopolítico por parte das empresas. Ao mostrar que os media estão entre esses fatores, esta dissertação contribui para uma melhor compreensão da perceção de risco, especialmente Risco Geopolítico, e da influência dos media que pode ser aplicada não só a empresas portuguesas, mas também a mercados internacionais. Além disso, a metodologia seguida constitui, por si só, um contributo relevante para a literatura, devido à utilização de LLMs para interpretar texto.

Palavras-Chave: Risco Geopolítico; Perceção de Risco Geopolítico; Perceção de Risco pelas Empresas; Cobertura dos Media; Framing dos Media; Empresas Portuguesas; BRGE2025

Table of Contents

1. Introduction	1
2. Literature Review	3
2.1. <i>First Strand: Measuring Geopolitical Risk</i>	3
2.2. <i>Second Strand: Behind Risk Perception</i>	7
2.3. <i>Research Niche</i>	11
3. Methodology	12
3.1. <i>BRGE2025</i>	12
3.1.1. Analysis	12
3.1.2. Calculations	14
3.2. <i>MW-GPR Index</i>	15
3.2.1. Data Extraction	16
3.2.2. Construction of the Database	20
3.2.3. Data Treatment	21
3.2.4. Calculations	23
3.3. <i>Combination of BRGE2025 and MW-GPR Index</i>	26
4. Results	27
4.1. <i>BRGE2025</i>	27
4.1.1. Short Term	27
4.1.2. Long Term	29
4.2. <i>MW-GPR Index</i>	31
4.2.1. Intermediate Results	32
4.2.2. MW-GPR Final Indicators' Results	35
4.3. <i>Combination of BRGE2025 and MW-GPR Index (Final Results)</i>	37
4.3.1. Short Term	37
4.3.2. Long Term	41
5. Limitations and Future Paths of Study	43
6. Conclusion	45
7. Appendix – Variables, Dimensions and Indicators Glossary	47
8. References	48
Annexes	I
<i>Annex I: Methodology followed by the BRGE2025</i>	I
<i>Annex II: Google Search Queries</i>	III
<i>Annex III: Extraction AI Prompt</i>	XIII
<i>Annex IV: Extracted HTML</i>	XVI
<i>Annex V: Opinion Piece Removal AI Prompt</i>	XVII
<i>Annex VI: Opinion Piece Removal Stata Validation Sample Prompt</i>	XIX
<i>Annex VII: Opinion Piece Removal Validation Sample Statistical Details</i>	XXVI
<i>Annex VIII: Relevancy Classification Python Prompt</i>	XXVII

<i>Annex IX: Relevancy Classification Stata Validation Sample Prompt Example – GPR Group 1</i>	<i>XXXIV</i>
<i>Annex X: Relevancy Classification Validation Sample Statistical Details.....</i>	<i>XXXVIII</i>
<i>Annex XI: Framing Classification AI Prompt.....</i>	<i>XL</i>
<i>Annex XII: Detailed Analysis of the Final Calculations of the MW-GPR Index.....</i>	<i>XLIV</i>
<i>Annex XIII: Initial MW-GPR Index Database</i>	<i>XLVII</i>
<i>Annex XIV: Opinion Piece Removal Testing Sample.....</i>	<i>XLVIII</i>
<i>Annex XV: MW-GPR Index Database After Removal of Opinion Pieces</i>	<i>XLIX</i>
<i>Annex XVI: Irrelevant Articles Removal Testing Samples.....</i>	<i>L</i>
<i>Annex XVII: Sensitivity Analysis I.....</i>	<i>LIII</i>
<i>Annex XVIII: Sensitivity Analysis II</i>	<i>LIX</i>

List of Images

Figure 1 - Ways to measure GPR perception, positioning the MW-GPR Index and BRGE2025 within the broader set of existing measurement instruments. Source: own elaboration.	4
--	---

List of Tables

Table 1 - Comparative overview of the main GPR barometers. Source: own elaboration.	6
Table 2 - Original BRGE2025 GPR groups and corresponding simplified English labels used in this study. Source: own elaboration.	13
Table 3 - Three-phase methodology used to construct the MW-GPR Index. Source: own elaboration.	15
Table 4 - Portuguese search keywords used to identify news articles for each GPR group. Source: own elaboration.	19
Table 5 - Coefficient values assigned to date, newspaper and headline framing for the calculation of article sentiment. Source: own elaboration.	25
Table 6 - Short-term BRGE2025 average perceived GPR scores, measured on a 1-5 scale, by GPR group, firm size and Openness Score (OS), with superscript numbers indicating each group's ranking within the respective subdivision. Source: own elaboration.	27
Table 7 - Long-term BRGE2025 average perceived GPR scores, measured on a 1-5 scale, by GPR group, firm size and Openness Score (OS), with superscript numbers indicating each group's ranking within the respective subdivision. Source: own elaboration.	29
Table 8 - Distribution and number of extracted articles by GPR group, keyword and Portuguese online newspaper. Source: own elaboration.	33
Table 9 - Number of articles classified as Concerning, Neutral or Dismissive based on headline framing. Source: own elaboration.	34
Table 10 - MW-GPR Index results by GPR group and indicator, measured in a scale of 1-5, with superscript numbers indicating each group's ranking within the respective indicator. Source: own elaboration.	35
Table 11 - Short-term correlation coefficients (-1 to 1) between MW-GPR Index indicators and BRGE2025 perceived GPR scores, by firm size and Openness Score (OS). Source: own elaboration.	37
Table 12 - Long-term correlation coefficients (-1 to 1) between MW-GPR Index indicators and BRGE2025 perceived GPR scores, by firm size and Openness Score (OS). Source: own elaboration.	41
Table 13 - Variables, Dimensions and Indicators Glossary. Source: own elaboration	47

1. Introduction

Currently, Geopolitical Risk (henceforth designated solely as GPR) is one of the most uncertain types of risk in the global economy. Every day, countries and most especially, companies are affected by wars, commercial conflicts, cyberattacks or migration crises that stem from an increasingly interconnected world. Current evidence suggests that firms' GPR perception has been rising sharply. This affects supply chains, access to foreign markets, production costs, investment decisions and, ultimately, the way firms operate and prepare for the future (Carney, 2016; Yu and Wang, 2023; Shen, 2025). However, despite its growing importance, the way companies perceive GPR and the channels through which these perceptions are formed, remain relatively understudied.

The literature suggests that risk perception doesn't always reflect objective exposure to risk, but a combination of psychological, social and informational factors. The media is believed to be among these factors and plays a particularly relevant part as it can influence not only the visibility of those risks, but also the framing of events, potentially shaping how individuals and organizations perceive uncertainty and potential threats. However, while literature has analyzed the relationship between media and risk perception, it mostly focused in areas such as health crises, environmental risks and financial uncertainty, leaving the specific case of GPR largely understudied.

This thesis aims to address that gap and assess whether the GPR that received more attention or stronger framing in Portuguese media are also the GPR that Portuguese companies perceived as the most relevant. Therefore, the main research question of this dissertation is: **what is the relationship between Portuguese media coverage of GPR and the way Portuguese companies perceive those same GPR?** Additionally, this study aims to answer other more specific questions that stem from the main question, such as: **“is corporate GPR perception more closely related to the Volume of news published about each risk, or to the Sentiment expressed in news headlines?”**; **“is the relationship stronger in the short or long term?”**; **“do company size and openness to foreign markets influence this relationship?”** and/or **“to what extent is Abnormal Sentiment associate with GPR perception among firms?”**.

To answer these questions, a thorough literature review was conducted, introducing the concepts of GPR, risk perception, media influence, cognitive heuristics and framing effects.

This first chapter will lay the foundation for the methodology that will then explain how this study builds its own media-based indicator of GPR, using articles published by Portuguese online newspapers. A central and innovative feature of this methodology is the use of LLMs to process and interpret textual data, particularly to transform extracted information, identify opinion pieces and classify the framing expressed in news headlines. These applications are complemented by Google-based data collection, Python-based classification, robustness tests and other computational tools. The resulting indicator is then compared with survey-based data on how Portuguese companies perceive those same risks.

Then, the results will be presented and discussed. As will be deeply analyzed below, the findings of this study suggest that the relationship between Portuguese media coverage and corporate GPR perception is more closely associated with how risks are framed than with how frequently they are reported. While Volume shows weak or predominantly negative correlations with firms' perceptions, Sentiment is positively associated with them, particularly among larger and more internationally open firms.

Finally, the main contributions and limitations of the study and future research pathways will be identified. By focusing on the unexplored relationship between media coverage and corporate GPR perception and focusing on the Portuguese case, this dissertation makes methodological (by connecting media-based indicators with firm-level survey evidence) and empirical (by providing context-specific evidence on Portuguese firms) contributions to the literature.

2. Literature Review

This literature review is organized into two complementary strands. The first examines the definition of GPR and explains how it can be measured, paying special attention to media-based indexes and survey-based barometers. It will answer questions like “*What is GPR?*” or “*How can GPR be measured?*”.

The second strand focuses on how risk perceptions are formed, emphasizing the role of cognitive heuristics, emotional responses, media exposure and framing effects. It will answer questions like “*What factors influences risk perception?*”, “*Is media one of those factors?*” and “*How does media influence GPR perception?*”.

2.1. First Strand: Measuring Geopolitical Risk

Geopolitical Risk can be defined as the combination of two major elements: Risk and Geopolitics.

Risk is the product of the possibility, probability or likelihood of the occurrence of a hazard by a society’s exposure and vulnerability to that hazard (Kasperson et al., 1988; Beck, 1992; Gill et al., 2014; Cambridge Dictionary 2025b).

Geopolitics is the science that studies the way the geographical conditions of a country influence its opportunities, constraints, politics and overall relationships with other countries (Scholvin, 2016; Caldara and Iacoviello, 2022; Cambridge Dictionary 2025a).

Therefore, GPR is the possibility of emergence, materialization and intensification of conflicts, such as wars, acts of terrorism and/or tensions between states and political actors, that have the potential to affect the peaceful course of international relations (Caldara and Iacoviello, 2022). From a firm-level perspective, however, GPR is not only an external phenomenon, but a strategic variable that directly affects the company’s performance, supply chains, market access and regulatory framing (Rice and Zegart, 2019).

Despite its relevance, GPR is inherently challenging to measure, as objective and universal quantification is neither available nor adequate for capturing its multidimensional nature (Caldara and Iacoviello, 2022; Bondarenko et al., 2024). Additionally, businesses often lack internal structures, analytical tools and complete or objective information regarding geopolitical developments required to systematically assess GPR (Rice and Zegart, 2019).

Thus, risk perception, i.e. the mental process by which individuals evaluate the likelihood of the risks to which they are exposed (Slovic, 1987; Wahlberg and Sjöberg, 2000; Pinto et

al., 2023; Jones, 2025), assumes a pivotal role in the understanding of GPR measurement. In fact, there is a broad consensus in literature that GPR must be measured indirectly, through perception, as it reflects the ways individuals interpret, anticipate and respond to geopolitical threats.

In this regard, perception-based instruments are particularly appropriate to measure GPR. Literature reveals two main paths: risk indexes and risk barometers. These two paths can be analyzed in the image below (Figure 1).

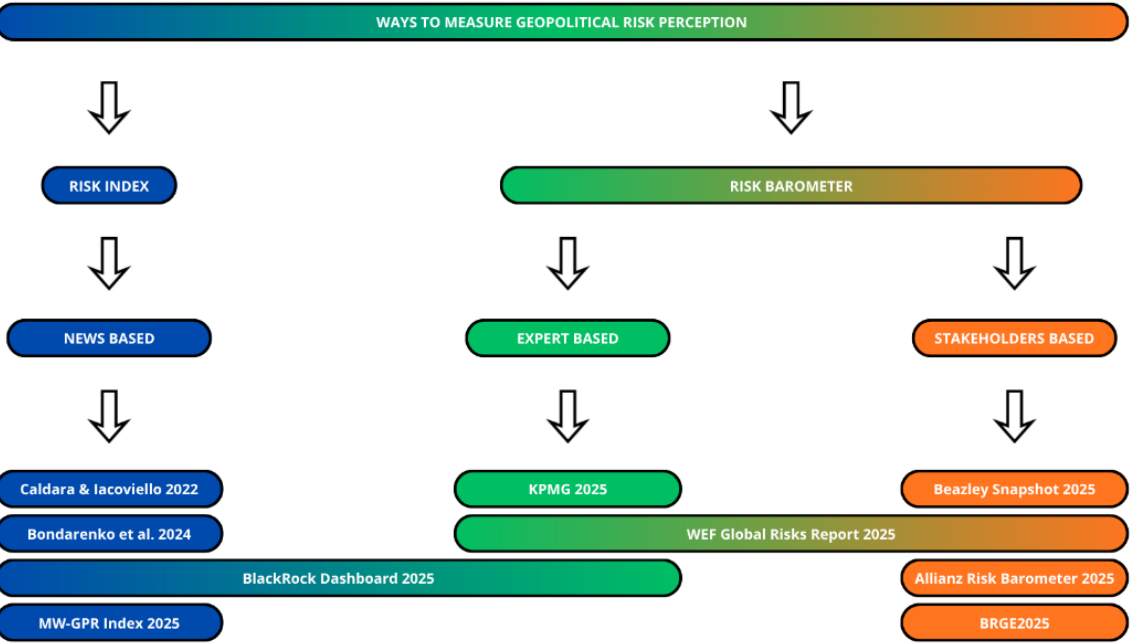


Figure 1 - Ways to measure GPR perception, positioning the MW-GPR Index and BRGE2025 within the broader set of existing measurement instruments. Source: own elaboration.

Risk indexes are usually based on news and media, following the seminal work of **Caldara and Iacoviello (2022)**. The authors created a GPR Index by examining the monthly percentage of published articles that discuss adverse geopolitical developments and associated risks from 1900 to the present day. This represents a sample of about 25 million articles, obtained by automated keyword-search on the digital archives of 10 leading English-language newspapers (*The Chicago Tribune, The Daily Telegraph, The Financial Times, The Globe and Mail, The Guardian, The Los Angeles Times, The New York Times, USA Today, The Wall Street Journal* and *The Washington Post*). The GPR Index is then calculated by dividing the number of articles discussing GPR by the total number of published articles (Caldara and Iacoviello, 2022). By

using media coverage as a proxy, this methodology captures fluctuations in GPR as reflected in the information environment. Variants of this methodology have since been applied to Russian and German language newspapers, which shows both robustness and wide adoption of this approach (Bondarenko et al., 2024).

However, as a media-based measure, the index shows patterns of news production and attention instead of how such risks are interpreted by economic agents. In other words, they capture *what is being discussed*, instead of *the perceived risk that stems from those discussions*.

To address this limitation, a second path relies on survey-based instruments, commonly referred to as risk barometers.

Risk barometers are structured survey instruments designed to capture the way individuals or organizations perceive and respond to different types of risk. Rooted in the broader literature on risk perception and survey-based measurement, barometers recognize that risk is a subjective evaluation shaped by expectations, experience and information. As such, they are particularly suited to contexts such as GPR, where objective measurement is limited and perceptions play a central role (Porto Business School, 2025). As seen in **Figure 1**, they can either be based on experts' or stakeholders' opinions.

To illustrate this diversity, **Table 1** presents a comparative overview of the main risk barometers that are relevant for the study of GPR and business risk perception. The table compares each instrument according to its approach, data source, purpose, time horizon and key feature, allowing a clearer understanding of how GPR perception can be measured in different institutional and business contexts.

This comparison shows that no single barometer captures GPR perception in a complete way. Expert-based instruments, such as KPMG, provide strategic interpretation, while hybrid models, such as BlackRock and WEF, combine qualitative judgement with media-based evidence or quantitative, respectively. Survey-based instruments, such as Beazley, Allianz and BRGE2025, are particularly relevant because they capture the perceptions of firms, executives or risk professionals more directly.

Additionally, it shows that there is not one right way to measure perceptions of GPR. Instead, different instruments capture different dimensions of risk perception, depending on their data source, methodological approach and target audience. However, none of these instruments fully understands how those perceptions are formed. For this reason, the next strand of the literature review focuses on the main factors that shape GPR perception.

Instrument	Approach	Data Source	Purpose	Time Horizon	Key Feature
KPMG	Qualitative, expert-based	Expert interviews and workshops	Identify GPR affecting sectors and provide strategic recommendations	Short term	Sector-specific assessment of GPR through expert judgement
BlackRock	Hybrid qualitative and quantitative	Media-based model and expert judgement	Monitor and assess the potential impact of GPR on global financial markets	Short term	Tailored to investors, combines GPR narratives with financial markets
Beazley	Quantitative, survey-based	Survey of international business leaders	Help firms build resilience	Short term	Focuses on how firms perceive uncertainty and resilience needs
Allianz	Quantitative, survey-based	Survey of risk management professionals from the insurance sector	Identify the most relevant business risks and strengthen firms' resilience	Short term	Tailored to the insurance sector, combines GPR with insurance risk exposure
BRGE	Quantitative, survey-based	Survey of Portuguese executives and business owners	Assess Portuguese firms' perceptions of GPR and provide mitigation tools	Short (1 year) and long (3 years) term	Provides country-specific evidence on how Portuguese firms perceive GPR
WEF	Hybrid survey-based and qualitative	Expert surveys, executive survey and workshops	Assess and anticipate global GPR, their probability and their impact	Short, medium and long term	Combines global comparability, multiple time horizons and cross-risk interconnection analysis

Table 1 - Comparative overview of the main GPR barometers. Source: own elaboration.

2.2. Second Strand: Behind Risk Perception

Having established that GPR must be measured through perception, it is necessary to understand how those perceptions are formed. While media-based indexes and barometers capture risk perception, they do not fully explain the cognitive, emotional and informational processes behind it. This second strand of the literature review addresses that issue.

Literature suggests that perceived risk is mainly composed of two dimensions: cognitive and emotional. The cognitive dimension contains knowledge, controllability and familiarity. It is usually associated with the likelihood of severity of the consequences, based on the available scientific information and on an objective appreciation. On the other hand, the emotional dimension is an intuitive response that includes dread and immediacy and expresses worry, anxiety and fear associated with the risk (Jones, 2025).

As a consequence, perceived risk doesn't always reflect actual risk with accuracy (Niu et al., 2020), partly because risk perception isn't formed unidirectionally (Kasperson et al., 1988) but, most importantly, due to cognitive biases and heuristics (Wahlberg and Sjöberg, 2000) that distort rational thinking and, consequently, risk perception (Jones, 2025). Therefore, we can conclude that all factors (both internal and external) that trigger these biases are sure to play a major role in the overall risk perception (Jones, 2025).

At firm level, GPR perception is also shaped by organizational priorities, managerial interpretation and structural characteristics. Firms may rely on familiar market logics, which can create blind spots and lead GPR to be underestimated or treated as a temporary rather than structural risk (Rice and Zegart, 2019; Rice, 2018). Evidence from SMEs further shows that risk perceptions vary with business situation, firm size and organizational structure, while remaining partly subjective and heterogeneous (Glowka et al., 2024).

Among external factors, literature highlights the intensity and proximity to geopolitical conflicts, the political and institutional context in which media outlets operate, government communication and censorship, diplomatic relations and economic exposure to affected regions (Bondarenko et al., 2024).

The literature also counts media¹ to be among these factors as, after all, most of the information we consume about risks is presented by the media (Singer and Endreny 1994; Niu et al., 2020; Jones, 2025). Although there isn't a broad consensus whether the effects of media coverage affect the decisions of government and businesses (Oliveira et al., 2025),

¹ According to Cunha and Peixinho (2020), media is not only the means used to transmit messages, but also the content of those means, the content industries and the technological devices used to spread the message.

literature suggests a strong connection between media coverage and perception of risk (Wahlberg and Sjöberg, 2000; Moussaïd, 2013; Jones, 2025).

However, not all risks seem to react in the same way to media exposure (Wahlberg and Sjöberg, 2000; Niu et al., 2020). In fact, literature believes that risk type may be a relevant moderating factor that affects the relationship between the media and the public's risk perception (Niu et al., 2020).

According to Wahlberg and Sjöberg (2000), media clearly influences general risk perception, i.e. people's judgments about risks at a societal or public level. However, personal risk perception, i.e. individuals' judgments about their own vulnerability to a risk, appears to be resistant to media and news influences. This is related to personal experience and/or to the way the public get their information (Wahlberg and Sjöberg, 2000). For instance, people who had access to direct information about a certain event were more likely to ignore the information transmitted by the media as they would find it inaccurate (Wahlberg and Sjöberg, 2000; Niu et al., 2020).

By definition, GPR affects not one individual but the entire society. Therefore, it is mostly considered as general risk (Caldara and Iacoviello, 2022). Literature indicates that this type of risk is mainly influenced by two specific components of media: volume of coverage and affective tone of the news (Wahlberg and Sjöberg, 2000; Lauriola et al., 2024).

Wahlberg and Sjöberg (2000)² examined the relationship between media coverage and risk perception without focusing on any specific type of risk and discovered that people's opinions are influenced more by the amount of media coverage than by its content. However, the same study concluded that it's probable that dramatic news highlight risk perception more than stories attached to the real facts. Additionally, they also concluded that the portrayal of an allegedly enraged public significantly increased their risk perception.

Building on these findings, the authors argue that the influence of media on risk perception is predominantly short-term, with its effects tending to fade over time. Importantly, they emphasize that the relationship between media and the public is not unidirectional, as audience reactions and demand for certain types of coverage shapes media output, which creates a reciprocal dynamic between news production and public perception. (Wahlberg and Sjöberg, 2000).

When considered collectively, these findings suggest that risk perception is shaped less

² Although this work predates much of the recent empirical literature on risk perception, its relevance remains undiminished, as it has since been reaffirmed and extended by recent studies.

by the substantive informational content of media coverage (*what is said*) than by form (*how it's said*) and frequency (*how many times it's said*), which can be explained through cognitive heuristics and biases, namely *availability* and *representativeness* heuristics and framing effects (Bondarenko et al., 2024; Lauriola et al., 2024).

The availability heuristic is a cognitive shortcut through which individuals assess the probability, frequency or causes of an event based on the ease with which instances of that event come to mind (Bazerman and Moore, 2013). It relies on factors such as *ease of recall*, i.e. vivid, recent or emotionally salient information is more readily remembered, and *retrievability*, which reflects how easily information can be accessed from memory. Although these factors are not biases in themselves, reliance on them can lead to systematic biases in judgment when easily recalled information is taken as representative of actual likelihoods. Literature categorized this heuristic as “media’s most fundamental way of altering people’s risk perception” (Wahlberg and Sjöberg, 2000, p. 44), as media exposure strongly conditions which events are cognitively accessible. This is directly related to the notion of risk perception and so it can be argued that the more volume of news a media outlet releases, the more available it is for the public and thus the more probable the public will perceive them.

The representativeness heuristic operates in a similar manner, influencing judgments by leading individuals to evaluate the probability of an event based on its perceived similarity to an existing mental category or prototype, rather than on statistical reasoning or base rates (Bazerman and Moore, 2013). This heuristic can produce biased judgments when surface characteristics are mistaken for genuine likelihood. Wahlberg and Sjöberg (2000) argue that media coverage plays a crucial role in shaping these initial representations by repeatedly associating risks with specific narratives or images. Once formed, these representations may bias subsequent interpretations of related events, reinforcing particular perceptions of risk.

Finally, framing effects interact with these heuristics by shaping how risk-related information is processed and interpreted in the first place. Different presentations of identical information, through variations in emphasis, language, narrative structure or emotional tone can trigger distinct cognitive and emotional responses, affecting judgments without altering the underlying facts (Wahlberg and Sjöberg, 2000; Bazerman and Moore, 2013).

Evidence from tourism studies also shows that differently framed news about political instability and terrorism can alter perceived destination risk, even when the underlying event remains constant (Kapuściński and Richards, 2016; Menon et al., 2025).

In the context of GPR, media framing determines not only which events receive atten-

tion but also how these events are interpreted in terms of potential consequences, urgency and severity (Bondarenko et al., 2024; Lauriola et al., 2024).

By emphasizing specific aspects of a risk, framing can attenuate or amplify the effects of the availability and representativeness heuristics, directing attention toward certain risks while downplaying others. Bondarenko et al. (2024) showed that perceived GPR varies systematically across media sources and national contexts, even when underlying geopolitical events are similar. Their findings suggest that differences in media agendas (strongly related to the extent of freedom of speech) emphasis and narrative focus play a crucial role in shaping how GPR are perceived locally. This reinforces the idea that perception is driven not only by events themselves but by how those events are framed and contextualized in the media.

This mechanism is consistent with framing effects, as repeated exposure to narratives emphasizing threat or uncertainty increases the perceived severity of GPR, even in the absence of objective changes in the international environment. Therefore, framing shapes not only which GPR becomes cognitively available, but also which emotional responses are associated with them, reinforcing heuristic-based judgments of likelihood and impact.

This theoretical relationship between framing, salience and emotion was empirically demonstrated in the context of GPR by Lauriola et al. (2024), albeit briefly. Focusing on Italian news articles covering the hypothesis of nuclear war, Lauriola et al. (2024) examined how both the volume and the emotional tone of news articles influence public perceptions of GPR and reported anxiety levels. Using a content analysis approach, the authors quantified media coverage by counting the number of articles published over time, while emotional tone was measured through sentiment analysis of the news content. Public perception of GPR and anxiety were then assessed using survey-based measures, which allowed the authors to examine how changes in the media relate to shifts in individual responses.

Their findings indicate that when both media coverage and emotional tone reached higher levels, public perception of GPR was heavily inflated and anxiety levels were also reportedly higher. However, when the media coverage dropped and the tone remained at the same level, public perception of GPR decreased and anxiety levels remained high (Lauriola et al., 2024). This led the researchers to suggest that public perception of GPR is connected with the sheer number of media articles released, while feelings of anxiety were related to the emotional tone of the news (Lauriola et al., 2024). However, as we learned above, through Jones' (2025) work, feelings of anxiety are a part of the public risk perception.

Taken together, these findings reinforce that media influence on GPR perception cannot

be reduced to a single dimension but instead reflects the combined effects of informational intensity and emotional framing.

2.3. Research Niche

The literature reviewed in this chapter can be organized into two complementary strands. The first focuses on the measurement of GPR perceptions through instruments such as risk barometers (Porto Business School, 2025). The second strand examines how risk perceptions are formed and influenced, emphasizing the role of media exposure, cognitive heuristics, emotional responses and framing effects (Wahlberg and Sjoberg, 2000; Niu et al., 2020; Lauriola et al., 2024; Jones, 2025).

However, with the exception of Lauriola et al. (2024), the literature has paid limited attention to the interaction between these two strands. In other words, existing research has generally examined either how GPR perceptions can be measured or how media coverage may shape risk perception, but has rarely connected media-based indicators with survey-based measures of corporate GPR perception.

This dissertation addresses that gap by examining whether the GPR categories perceived as most relevant by Portuguese firms, as measured by the Barómetro do Risco Geopolítico para Empresas (BRGE2025), are aligned with those most intensively covered and framed by Portuguese online media in the year preceding the survey, as measured by the Media-Weighted Geopolitical Risk Index (MW-GPR Index), an indicator created specifically for this study. By combining media-based indicators of risk coverage and tone with survey-based measures of firms' risk perceptions, the dissertation creates a unified empirical framework linking media dynamics and barometer-based perceptions. It also contributes context-specific evidence by focusing on Portugal during a period of heightened geopolitical uncertainty.

3. Methodology

The methodology followed stands over three axis: BRGE2025, MW-GPR Index and the combination of both. Therefore, this chapter will be divided into three main subsections.

The first subsection analyzes the BRGE2025 and justifies its selection over the other stakeholder-based barometers mentioned during the Literature Review. The necessary calculations to process the raw data of the barometer will also be explained.

Then, the second subsection focuses on the construction of the MW-GPR Index, from the decision to create a database from scratch to the assumptions made to extract the data and to the calculations used to transform the raw data into a stable index.

Finally, the third subsection studies the comparison between the GPR category rankings of the several dimensions of BRGE2025 and MW-GPR Index.

3.1. BRGE2025

As explained above, this subsection will be dedicated to the BRGE2025. First, the barometer will be presented, alongside its main characteristics, and analyzed. Then, the calculations needed to create the rankings will be presented and explained.

3.1.1. Analysis

BRGE2025 is a risk barometer, the first systematic project in Portugal that aims to evaluate the way Portuguese companies perceive GPR, while studying its impact on Portuguese corporate activity. Addressed to executives of businesses based in Portugal, it combines firm-level perspective, standardized company information, differentiated time horizons (short [1 year] and medium/long term [3 years³]) and a standardized survey structure aimed at capturing perceived relevance rather than expert assessments⁴.

Unlike investor-focused dashboards or globally aggregated surveys, BRGE2025 is explicitly designed to measure how firms perceive GPR within the clearly defined national context of Portugal. This factor cannot be overstated, as according to the findings of Bondarenko et al. (2024), the specific national business environment highly influences the way GPR is perceived. Therefore, as this study aims to focus solely on the specific case of

³ The decision to limit the medium/long term to only 3 years was made, not only due to the current high volatility of the international geopolitical context, but also to discard the impact that excessively long projections could have on the credibility of the answers (Porto Business School, 2025).

⁴ The full methodology followed by the BRGE2025 can be found in Annex I.

Portugal, it's appropriate to include a risk barometer focused on Portuguese companies.

Furthermore, by collecting additional information, such as firm size, internationalization activities and sector of activity, BRGE2025 allows this dissertation to analyze different types of companies, enriching its analysis.

Finally, by intentionally simplifying its survey, using Likert-scale⁵ questions to assess risks and yes/no questions to verify firm-level characteristics, BRGE2025 shows understanding of the current business landscape of Portugal, while guaranteeing results reliability and reducing response bias. This dimension is extremely important, as it ensures that Portugal's business reality is accurately represented, thereby enhancing the reliability of this study.

This planned simplification is also present in the decision to create 12 categories of GPR. By selecting the 12 groups of GPR considered to be more relevant, the survey reduces the risk of respondents overlooking a category and prevents the same risk from being reported under different labels. Below (**Table 2**) are presented the 12 groups of GPR selected by the BRGE2025, as well as the label by which they will be referenced in this study.

Group Of GPR in the BRGE2025	Label
1 - Novos conflitos globais ou descontrolo dos atuais	Global Armed Conflict
2 - Novos conflitos na Europa ou descontrolo dos atuais	Armed Conflict in Europe
3 - Explosão nuclear ou libertação de agentes biológicos ou químicos	Nuclear, Biological and Chemical Threats
4 - Ciberataques de grande dimensão a infraestruturas críticas ou empresas	Cyberattacks
5 - Migrações em massa usadas como arma política	Weaponized Migration
6 - Radicalização social e política, greves e motins civis violentos	Civil and Political Unrest
7 - Choques energéticos ligados ao petróleo e gás natural	Energy Shocks
8 - Crise financeira grave	Financial Crisis
9 - Disrupção das cadeias de abastecimento globais	Supply Chain Disruption
10 - Conflitos comerciais de elevada intensidade (EUA, China, UE)	Trade Wars
11 - Desinformação gerada por Inteligência Artificial	AI Disinformation
12 - Catástrofes climáticas extremas	Climate Disasters

Table 2 - Original BRGE2025 GPR groups and corresponding simplified English labels used in this study. Source: own elaboration.

⁵ Likert-scale is a survey tool used to measure respondent's level of agreement with a statement, usually in a 1-5 or 1-7 scale.

These features make BRGE2025 a highly appropriate and contextually grounded instrument for this study, that ensures the analysis is both empirically robust and closely aligned with the Portuguese business reality. However, for the study to better use these characteristics, there were some restrictions that had to be applied to the MW-GPR Index before and during its construction. These will be thoroughly explained in the following sections.

3.1.2. Calculations

The approach initially followed by the barometer aimed at a media-friendly, commercial reading of the barometer's results. That approach, while extremely effective for the goals of the barometer, proved to be restrictive to the goals of this study, as it disregarded valuable information in the search for simpler conclusions. Therefore, a simple average was computed for each GPR group. This method was chosen because all responses were already expressed on the same 1 to 5 scale, meaning that each observation could be treated equally without requiring further transformation. Moreover, in the absence of a clear methodological basis for assigning different weights to specific firms or responses, the average provides a transparent, neutral and easily comparable measure of perceived risk across GPR groups.

To conduct a deeper analysis, the data was filtered according to time horizon, firm size and openness score. This allowed the analysis to account for variations across different time spans and company profiles.

As the surveys from BRGE2025 separate the answers into short- and long-term, no external calculations had to be made. Simply filtering the answers according to their original time horizon made it possible to compute separate averages for each GPR group in the short and long term, preserving the structure of the BRGE2025 while allowing a direct comparison with the MW-GPR Index.

Regarding the size of the company, the surveys from BRGE2025 already separated the companies into “Micro Empresa” (Micro Firms), “PME” (SMEs) and “Grande Empresa” (Large Firms). Therefore, no external calculations were necessary. Simply filtering the answers according to these predefined categories made it possible to compute the average score for each GPR group within each company-size segment, allowing the comparison to reflect potential differences in risk perception between smaller and larger firms.

Additionally, the degree of openness of each company, i.e., how open to foreign markets the companies are, was also considered. Despite asking whether the company exports, im-

ports and/or has FDI abroad, the BRGE2025 doesn't compute a direct openness score. Therefore, to filter this dimension, an openness score was created by assigning one point to each form of international exposure reported by the company. As a result, companies with more international activities were classified as having a higher degree of openness (OS), allowing the analysis to compare risk perceptions according to their level of exposure to external markets. Therefore:

- **OS = 0:** if a company doesn't export, doesn't import and doesn't have FDI abroad;
- **OS = 1:** if a company either exports, imports or has FDI abroad;
- **OS = 2:** if a company does two of the three activities (exports, imports, FDI abroad);
- **OS = 3:** if a company exports, imports and has FDI abroad.

3.2. MW-GPR Index

As explained above, this subsection will be dedicated to the construction of the MW-GPR Index. This was divided into three phases, as the following table (**Table 3**) explains:

1. Data Extraction	This phase defined the scope of the analysis, including the study period, selected Portuguese online newspapers and keywords associated with each GPR category. Targeted Google searches were then used to identify relevant articles and construct the initial database. LLMs were used to ease the search.
2. Data Treatment	This phase cleaned and classified the collected data. Opinion pieces and articles unrelated to GPR were removed, while the remaining headlines were classified according to their framing. LLMs supported the removal of opinion pieces (backed by robustness tests) and the framing classification (fully relying on AI to interpret and assess the meaning and tone conveyed by the text), whereas a Python-based model (backed by robustness tests) was used to assess article relevance.
3. Calculations	This phase converted the treated data into the final MW-GPR indicators. Volume, Sentiment, Composite and Abnormal Sentiment scores were calculated and standardized on a 1-5 scale.

Table 3 - Three-phase methodology used to construct the MW-GPR Index. Source: own elaboration.

This approach represents a coherent and well-justified methodological framework, combining the rigor of existing literature with the flexibility required to best use the BRGE2025. As such, it ensures that the MW-GPR Index is methodologically sound and well-suited to the objectives of this study⁶.

3.2.1. Data Extraction

Literature provides two paths to create a media-based index. A study can either create a database from scratch by querying Google or any other search engine (as done by Lauriola et al., 2024) or access a pre-existing database like Factiva (in the manner of Bondarenko et al., 2024) or The GDELT Project.

Both paths were considered, but the decision fell upon the former, due to the increased flexibility it allows. Furthermore, pre-existing databases are subject to access constraints and limited availability, especially as the objective of the study is narrow in scope and extensive and specific in time. Additionally, creating a database from scratch allowed this dissertation to shape the MW-GPR to the characteristics of the BRGE2025, as it permitted the MW-GPR Index to be much more tailored to the characteristics of the barometer than a pre-existing database ever could.

Therefore, the process followed a methodology similar to the one applied by Lauriola et al., (2024), where the process for extracting news content followed a structured keyword-based approach combined with web scraping techniques. Specifically, these authors used targeted search queries to retrieve news headlines from major online newspapers. Similarly, this study adopts a keyword-based strategy, relying on Google searches to identify relevant news articles from different media outlets.

To conduct the Google searches, some assumptions had to be made, namely regarding the Time Period on which the articles were published, the News Outlets from where the articles were extracted and the Keywords used for the said extraction.

⁶ It seems essential to mention that the work of Caldara and Iacoviello (2022) was not followed in any of these steps, due to differences in scope, data structure and research objectives. Despite being the most cited GPR media-based index, their method consisted of developing a large-scale, dictionary-based methodology suited for analyzing millions of articles over a long historical period (since 1900), an approach that was not appropriate for this study.

Time Period According to Wahlberg and Sjoberg (2000), it's not certain how long news articles influence people's perception of risk. However, as we saw above, literature suggests that the effect is much more intense in the short term than in the long term. Therefore, an assumption was made, and it was decided to analyze the one-year period prior to the completion of the surveys of BRGE2025. Thus, the time period from which news articles were collected was from 08/02/2024 to 08/02/2025⁷.

News Outlets Regarding the News Outlets on which the articles were published, the choice fell upon the online edition of 5 Portuguese newspapers: *Expresso*, *Jornal de Negócios*, *Público*, *Observador* and *Diário de Notícias*.

Firstly, the decision to include only Portuguese newspapers was done because a country's local press provides a more detailed narrative of national events (Andres-Escayola et al., 2022), which is especially relevant when the covered topic is GPR (Bondarenko et al., 2024). In fact, Bondarenko et al., 2024 discovered that the language on which the article is published highly influences the way GPR is perceived. Therefore, the choice fell upon Portuguese media only.

Then, the choice to consider solely the online edition of newspapers was based on their interactivity and immediacy and for providing quick access, constant updates and greater ease of consultation (Cunha and Peixinho, 2020). Additionally, no tabloids were included, following the recommendations of Andres-Escayola et al. (2022).

Finally, broad press coverage is essential when constructing textual indicators, as using a single outlet may lead to biased results (Andres-Escayola et al., 2022). Thus, this study relies on multiple newspaper sources. After thorough consideration and analysis of the current landscape of Portuguese media, the five newspapers were selected. The reasons for the selection of each outlet were as follows:

- **Expresso:** due to its wide reach and strong reputation within the Portuguese media landscape. As one of the most established and influential newspapers in Portugal, it provides comprehensive coverage of political and economic developments, making it a relevant source for capturing media narratives related to GPR.
- **Jornal de Negócios:** for its specialized focus on economic and financial topics. Its coverage of business, markets and macroeconomic developments makes it particu-

⁷ The last day in which the surveys of the BRGE2025 were conducted was 08 of February of 2025, according to the Methodological Appendix of the referred barometer, which can be found in **Annex I** of this dissertation.

larly relevant for identifying news related to economic dimensions of GPR, complementing the more generalist outlets included in the sample.

- **Público:** due to its broad readership and recognized credibility as a leading national newspaper. Known for its in-depth reporting and editorial rigor, it contributes to ensuring that the sample reflects high-quality and widely consumed media content.
- **Observador:** as a native digital newspaper, reflecting the growing importance of online media in shaping public discourse. Its digital-first model allows for rapid publication and continuous updates, making it particularly suitable for capturing timely reactions to geopolitical developments.
- **Diário de Notícias:** due to its institutional relevance and long-standing presence in the Portuguese media landscape. As one of the oldest newspapers in Portugal, it provides a more traditional and institutional perspective, contributing to a balanced representation of media narratives.

Keywords The selection of keywords was designed to be intentionally broad, aiming to capture the largest possible set of articles and minimize the risk of excluding pertinent content. Rather than relying on a predefined dictionary, the keywords were defined based on an understanding of the Portuguese media landscape and the typical language used in news. This approach ensured both flexibility and contextual relevance in the data extraction process.

With the exception of a small number of technical terms, all keywords were specified in Portuguese and, in some cases, were constructed using Boolean operators (e.g., AND, OR) to improve search accuracy. For each Group of GPR⁸, three complementary keywords were selected in order to capture different dimensions of the same risk. The keywords used can be seen in **Table 4**.

⁸ These Groups of GPR were the same used in the BRGE2025.

Group of GPR	Keyword	
1 - Global Armed Conflict	1	guerra AND militar AND internacional
	2	armamento AND militar
	3	NATO
2 - Armed Conflict in Europe	1	guerra AND militar AND (Ucrânia OR Europa)
	2	armamento AND militar AND (Europa OR Ucrânia)
	3	NATO AND (Ucrânia OR Europa)
3 - Nuclear, Biological and Chemical Threats	1	explosão AND nuclear
	2	arma AND (biológico OR químico)
	3	ameaça AND guerra AND nuclear
4 - Cyberattacks	1	ciberataque OR “ataque informático”
	2	ataque AND digital AND (infraestrutura OR empresa)
	3	Cibersegurança
5 - Weaponized Migration	1	migração AND segurança
	2	fronteira AND migrante
	3	refugiado AND crise
6 - Civil and Political Unrest	1	manifestação OR motim
	2	Greve
	3	radicalização AND (social OR política)
7 - Energy Shocks	1	abastecimento AND energia
	2	crise AND energética
	3	choque AND (petróleo OR gás)
8 - Financial Crisis	1	crise AND financeira
	2	Recessão
	3	instabilidade AND financeira
9 - Supply Chain Disruption	1	“cadeia de abastecimento” OR “cadeias de abastecimento”
	2	crise AND (logística OR abastecimento)
	3	"Canal de Suez" OR "Canal do Panamá" OR “Estreito de Ormuz”
10 - Trade Wars	1	“guerra comercial”
	2	tarifa AND internacional
	3	conflito AND comércio
11 - AI Disinformation	1	deepfake
	2	desinformação AND (“Inteligência Artificial” OR AI OR IA)
	3	manipulação AND informação
12 - Climate Disasters	1	clima AND (catástrofe OR extremo)
	2	“incêndio florestal” OR “incêndios florestais”
	3	“alteração climática” OR “alterações climáticas”

Table 4 - Portuguese search keywords used to identify news articles for each GPR group. Source: own elaboration.

3.2.2. Construction of the Database

Once the assumptions were established, the construction of the database was initiated. The specific process was as follows:

1. **Google Search:** a query in Google was made following the type⁹ “site:*web-site_name* *keyword* after:2024-02-08 before:2025-02-08”. All of the Google queries can be found in Annex II.
2. **Extraction:** to extract the links and the dates in the most efficient way, the HTML code was pasted in a Word file¹⁰ with the name “*GPR_Group_Code*. *Newspaper_Code*. *Keyword_Code*”¹¹;
3. **Transformation:** after having a Word file with the HTML containing all of the relevant information, there was a need to transform that HTML into a workable excel file. To achieve this, the Word file was uploaded to a LLM (Copilot or Chat GPT) with the prompt that can be found in **Annex III**. The excel was then downloaded and slightly treated before the Date and Link columns were copied into the definitive excel file on which the initial database was stored.

Although the definitive process was rather simple to execute, there were some difficulties along the way that should be mentioned to future researchers exploring this path:

- **AI Limitations:** could not directly conduct or automate the search and retrieval of news articles as initially intended. This led to the much more manual process described above and reduced the level of automation. On the other hand, the new method became much more reproducible;
- **Google Limitations:** the use of Google Search as the primary data source also introduced technical limitations. In particular, restrictions on web scraping led to partial blocking of HTML content, which hindered the automated extraction of search results. As a result, the data collection process was significantly delayed, requiring additional adjustments and workarounds to ensure the completeness of the dataset.

⁹ e.g., site:expresso.pt/ guerra AND militar AND internacional after:2024-02-08 before:2025-02-08

¹⁰ All of the extracted HTML are in the Google Drive whose link can be found in Annex IV.

¹¹ e.g., 1.X.1

3.2.3. Data Treatment

After extracting the links and respective headlines, there was the need to treat the data. This step followed a three-phased approach: removal of the opinion pieces, removal of the articles that weren't related to GPR and classification of each headline according to framing.

Removal of Opinion Pieces The first step was to remove the opinion pieces from the sample, following in the recommendations of Althaus et al. (2001). This restriction ensures greater consistency and comparability across observations, as such formats reflect subjective interpretations rather than standardized news reporting, potentially introducing bias into the analysis (Althaus et al., 2001). This point is particularly relevant for this study, as the MW-GPR Index aims to measure only volume and tone, not subjective interpretations or commentary. Those dimensions are captured by the BRGE2025 and thus, including opinion pieces would risk endogeneity in the final correlation.

To achieve this, the database was uploaded to a LLM, which classified each article as “Opinion Piece” or “Not Opinion Piece”, based on the information available on provided link. This step represents an important methodological feature of the study, as the LLM was used as an interpretative tool capable of assessing the nature and purpose of each article instead of text extraction or mechanical processing. The prompt that was used to complete this step can be found in **Annex V**.

However, given the interpretative nature of this task, the AI-generated classification was not accepted without validation. Afterwards, robustness tests were run on a random and heterogeneous (including articles from different newspapers, groups of GPR and publishing dates) training sample, created with Stata. The prompt used to create the sample can be seen in **Annex VI**.

This sample was then manually classified and compared with the LLM output in order to assess the accuracy and reliability of the AI classification. By combining LLM-based interpretation with manual validation, this methodology reduces the risk of relying on an unverified black-box classification process and strengthens the credibility of the final dataset. The statistical details of the sample can be found in **Annex VII**.

Removal of Irrelevant Articles The next step consisted of excluding articles that, despite containing the selected keywords, were not related to GPR and are, therefore, irrelevant to the sample. Considering that the keywords were deliberately broad, it's expected for the da-

tabase to include a lot of these irrelevant articles. As one of the key study variables of this dissertation is volume of articles published per Group of GPR, keeping them would seriously damage the results.

Lauriola's et al. (2024) method relied on dictionary-based filtering strategies, combining co-occurrence rules and exclusion terms to reduce misclassification and improve the accuracy of the selected content. Their study, however, focused on a single geopolitical event (the war in Ukraine). This study adopts a broader scope, and so, following this approach would require the construction of a separate dictionary for each of the 12 Groups of GPR considered in BRGE2025, which would be excessively time-consuming and complex. Therefore, this study opted for a more practical approach: a Python-based model.

Before developing it, however, a few different, more practical methods were tried. Following the framework of the previous chapter, the first attempt consisted of an AI LLM prompt. Nevertheless, this approach was not effective as the used LLM didn't have enough consistency in the classification, as proved by the Robustness Tests conducted posteriorly. The second attempt consisted in the training of a Knn model, which also proved ineffective.

Finally, a Python-based model was developed to more efficiently and flexibly identify articles that were substantively related to GPR. An example of the prompt used to complete this step can be found in **Annex VIII**.

Afterwards, robustness tests were run on a random and heterogeneous training sample (including articles from different newspapers and publishing dates), created with Stata. An example of the prompt used to create the samples can be seen in **Annex IX**.

Since the number of articles varied considerably across GPR groups, each category was represented by a different number of observations in order to ensure adequate coverage of the dataset. These samples were then manually classified and compared with the model's output to assess the accuracy of the classification. The details of the samples are as shown in **Annex X**.

Framing Classification Finally, the third step of this phase consisted of classifying each headline according to the type of framing. This step is extremely important, as this dissertation aims to measure not only the volume, but also the tone and sentiment expressed by the news.

Following the same procedure used for the removal of opinion pieces, headline framing

was classified using an LLM. Each headline was classified as “Concerning” (if the headline actively increases the perception of GPR), “Neutral” (if the headline is neutral or doesn’t influence the perception of GPR) or “Dismissive” (if the headline actively decreases the perception of GPR). This step is one of the main methodological contributions of the study, as the LLM was used not merely as an automation tool, but as an interpretative instrument capable of assessing meaning, tone and framing in short textual units. Unlike dictionary-based sentiment methods, which rely mainly on the presence of predefined words, this approach allows the classification to account for context, nuance and the way geopolitical risks are presented to readers.

The decision to classify only headlines rather than full articles is grounded in the fact that headlines often constitute the primary point of interaction between readers and news content. In many cases, readers engage only with the headline, using it as a heuristic to interpret and evaluate the underlying information. As a result, they can function as cognitive shortcuts through which individuals interpret relevance, urgency and severity of a given risk.

Additionally, focusing on headlines enhances the efficiency and consistency of the classification process, while still preserving the key informational signals relevant for this analysis. The prompt used to complete it can be found in **Annex XI**.

Unlike the classification of Opinion Pieces, no robustness tests were run in this classification. This was because this classification is entirely subjective and inherently dependent on interpretative judgment, making formal validation less reliable and straightforward.

Lauriola et al. (2024), who also aimed to measure sentiment, used a LIWC (Linguistic Inquiry and Word Count) paired with an Italian emotional dictionary (ItEM) that allowed to compute the frequency of words associated with emotions. However, this approach was not followed as an LLM has the unique ability to understand context, nuance and modal expressions. On the other hand, this comprehension can also add subjectivity and reduce reproducibility. This trade-off was acknowledged and found acceptable.

3.2.4. Calculations

Finally, the definitive values of the MW-GPR Index were computed, so that they can be compared with the results of BRGE2025 in the last stage of the methodology (3.3. Combination of BRGE2025 and MW-GPR Index). These calculations will be analyzed and explained in the following pages. A deeper, more detailed and technical analysis can be found

in **Annex XII**. Since the BRGE2025 results were all presented on a Likert-Scale (1 to 5), the Index results were also converted to this same scale to ensure a more consistent comparison. To be able to conduct a deeper analysis, the MW-GPR Index considers **Volume** (number of news' articles published) and **Sentiment** (emotional tone used in the headline of the news' articles), both in a **Composite** way (a weighted average of the two components) and separately, using the same approach as Lauriola et al. (2024). Additionally, **Abnormal Sentiment** (the unexpected variation in media sentiment in the most recent period of time) was also computed, adapting the method of Venade and Grilo (2023) to better fit this study.

Indicator 1: Volume Computed by standardizing the number of articles published for each GPR group using min-max normalization, i.e., comparing the number of articles in a given group with the minimum and maximum number of articles observed across all groups. The result was then converted into a 1-5 scale, where 1 corresponds to the least covered GPR group and 5 corresponds to the most covered GPR group¹². Therefore, higher values indicate greater relative media attention. The formula used is presented below, where i represents the Group of GPR and ranges from 1 to 12.

$$Volume_i = \frac{Number\ of\ articles_i - Minimum\ Number\ of\ articles}{Maximum\ Number\ of\ articles - Minimum\ Number\ of\ articles} \times 4 + 1$$

Indicator 2: Sentiment To measure the **Sentiment**, a coefficient-based approach was followed. This approach was chosen because it allows the sentiment to be measured in a systematic and comparable way that translates the main factors that influence it (the date each article was published (Date)¹³, the outlet where the article was published (Outlet)¹⁴ and the affective tone expressed by the headline (Framing)) into quantifiable values. Thus, the Sentiment of each article was computed by multiplying the three coefficients, as presented in the following formula, where k represents the articles. The values the coefficients could take, which were assumed and later validated by a sensitivity analysis, can be seen in the table below (Table 5).

¹² This method assumes that the sample of relevant articles used in this study includes all articles related to GPR published in Portugal during the selected period. This limitation was acknowledged and found acceptable.

¹³ According to Wahlberg and Sjoberg (2000), the influence of media coverage decreases over time.

¹⁴ According to Jones (2025), the source's trustworthiness positively influences the perception of GPR.

$$Sentiment_k = Date_k \times Outlet_k \times Framing_k$$

Dimension	Interval/Cathegory	Coefficient
Date	From 08 of November of 2024 to 08 of February of 2025	1
	From 08 of August of 2024 to 07 of November of 2024	0.75
	From 08 of May of 2024 to 07 of August of 2024	0.5
	From 08 of February of 2024 to 07 of May of 2024	0.25
Newspaper	Expresso	1.2
	Público	1.1
	Negócios	1.0
	Observador	0.9
	Diário de Notícias	0.8
Framing	Concerning	1.5
	Neutral	1.0
	Dismissive	0.5

Table 5 - Coefficient values assigned to date, newspaper and headline framing for the calculation of article sentiment. Source: own elaboration.

After computing the sentiment expressed by each article, it was necessary to obtain an overall sentiment score for each GPR group. To do so, the average sentiment of all articles belonging to each group was calculated. These average values were then standardized using min-max normalization. The result was then converted into a 1-5 scale, where 1 corresponds to the group with the lowest sentiment score and 5 to the group with the highest sentiment score. Thus, higher values indicate stronger relative sentiment. The formula used was as follows, where i represents the Group of GPR:

$$Sentiment_i = \frac{Sentiment_i - Minimum\ Sentiment}{Maximum\ Sentiment - Mininum\ Sentiment} \times 4 + 1$$

Indicator 3: Composite Achieved by computing a weighted average between the Sentiment and the Volume, as shown in the formula below (i represents the Group of GPR and ranges from 1 to 12). The weights used in this average were an assumption. However, to better understand their accuracy, a sensitivity analysis was conducted.

$$Composite_i = 0,5 \times Sentiment_i + 0,5 \times Volume_i$$

Indicator 4: Abnormal Sentiment Following Venade and Grilo (2023), Abnormal Sentiment captures the unexpected component of media sentiment by measuring the difference between observed sentiment and an expected benchmark.

Although this dissertation followed in the footsteps of Venade and Grilo (2023), some adjustments were made. Firstly, instead of using Sentiment alone, Abnormal Sentiment is calculated using the Composite variable, since literature suggests that media relevance depends both on the tone of coverage and on the volume of coverage. Then, as no external benchmark is available, the expected Composite value for the most recent quarter is estimated through a linear regression based on the first three quarters.

Abnormal Sentiment is then calculated as the Difference between the observed and expected Composite values and converted into a 1-5 scale using min-max normalization, as explained in the following formula, where i represents the GPR Groups.

$$Abnormal\ Sentiment_i = \frac{Difference_i - Minimum\ Difference}{Maximum\ Difference - Minimum\ Difference} \times 4 + 1$$

3.3. Combination of BRGE2025 and MW-GPR Index

Finally, the results obtained from the MW-GPR Index and the BRGE2025 were converted into rankings within the subdivisions analyzed above. The Spearman rank correlation coefficient was then applied between each MW-GPR Index indicator¹⁵ and each BRGE2025 dimension¹⁶. This method was chosen because the analysis focuses on the relative ordering of GPR groups, rather than on the exact numerical distance between their scores. It also doesn't require a linear relationship between variables or a normal distribution.

This procedure produced the final results of the study: a correlation coefficient ranging from -1 to 1, reflecting the relationship between the two rankings being compared. A value close to 1 indicates a very strong direct relationship, meaning that the two rankings are highly similar. A value close to -1 indicates a very strong inverse relationship, meaning that GPR groups ranked higher in one indicator tend to be ranked lower in the other. Values close to 0 indicate a weak or non-existent relationship between the rankings.

¹⁵ Volume, Sentiment, Composite and Abnormal Sentiment.

¹⁶ Short-term, Long-term, Micro Firms, SMEs, Large Firms, Openness Score = 0, Openness Score = 1, Openness Score = 2 and Openness Score = 3 for the BRGE2025.

4. Results

In the previous chapter, the Methodology followed in this dissertation was presented using a process point of view, presenting the different stages one after the other according to how they were performed. As a consequence, intermediate stages also produced relevant results that influenced the final results and must be mentioned in this Results chapter as well. Thus, this chapter will follow the same structure of the Methodology, i.e., firstly, the results of the BRGE2025 will be presented; then, the results of each stage of the construction of the MW-GPR Index will be presented, in the same order as the process shown in the Methodology; finally, the combination of both results will be presented, discussed and analyzed.

4.1. BRGE2025

After finalizing the computation of averages of each subdivision, the results of each Group of GPR were ranked within their subdivision and compared. These rankings (for the short and long term) will now be presented and analyzed.

4.1.1. Short Term

The table below (**Table 6**) reports the results of the BRGE2025 for the short term¹⁷.

Group Of GPR	Short Term	Micro Firms	SMEs	Large Firms	OS = 0	OS = 1	OS = 2	OS = 3
1 - Global Armed Conflict	3,625 ⁽⁵⁾	3,476 ⁽⁵⁾	3,681 ⁽⁵⁾	3,547 ⁽⁶⁾	3,571 ⁽⁴⁾	3,488 ⁽⁸⁾	3,458 ⁽⁸⁾	3,724 ⁽²⁾
2 - Armed Conflict in Europe	3,696 ⁽⁴⁾	3,643 ⁽³⁾	3,596 ⁽⁷⁾	3,679 ⁽⁴⁾	3,551 ⁽⁵⁾	3,659 ⁽⁵⁾	3,708 ⁽³⁾	3,655 ⁽⁴⁾
3 - Nuclear, Biological and Chemical Threats	3,262 ⁽¹¹⁾	3,171 ⁽¹²⁾	3,064 ⁽¹²⁾	3,415 ⁽⁸⁾	3,102 ⁽¹¹⁾	3,350 ⁽¹¹⁾	3,125 ⁽¹²⁾	3,276 ⁽¹⁰⁾
4 – Cyberattacks	3,875 ⁽²⁾	3,762 ⁽²⁾	3,936 ⁽²⁾	3,717 ⁽³⁾	3,898 ⁽¹⁾	3,829 ⁽²⁾	3,500 ⁽⁷⁾	3,793 ⁽¹⁾
5 - Weaponized Migration	3,109 ⁽¹²⁾	3,195 ⁽¹⁰⁾	3,085 ⁽¹¹⁾	2,906 ⁽¹²⁾	3,061 ⁽¹²⁾	3,000 ⁽¹²⁾	3,217 ⁽¹¹⁾	2,931 ⁽¹²⁾
6 - Civil and Political Unrest	3,359 ⁽¹⁰⁾	3,238 ⁽⁹⁾	3,404 ⁽¹⁰⁾	3,415 ⁽⁸⁾	3,245 ⁽⁹⁾	3,488 ⁽⁸⁾	3,333 ⁽⁹⁾	3,310 ⁽⁹⁾
7 - Energy Shocks	3,440 ⁽⁹⁾	3,310 ⁽⁸⁾	3,511 ⁽⁸⁾	3,321 ⁽¹¹⁾	3,449 ⁽⁷⁾	3,488 ⁽⁸⁾	3,292 ⁽¹⁰⁾	3,138 ⁽¹¹⁾
8 - Financial Crisis	3,913 ⁽¹⁾	3,952 ⁽¹⁾	3,872 ⁽³⁾	3,904 ⁽¹⁾	3,878 ⁽²⁾	4,049 ⁽¹⁾	3,958 ⁽¹⁾	3,643 ⁽⁵⁾
9 - Supply Chain Disruption	3,609 ⁽⁶⁾	3,381 ⁽⁷⁾	3,511 ⁽⁸⁾	3,736 ⁽²⁾	3,347 ⁽⁸⁾	3,756 ⁽⁴⁾	3,583 ⁽⁶⁾	3,552 ⁽⁶⁾
10 - Trade Wars	3,783 ⁽³⁾	3,643 ⁽³⁾	3,957 ⁽¹⁾	3,642 ⁽⁵⁾	3,673 ⁽³⁾	3,780 ⁽³⁾	3,833 ⁽²⁾	3,690 ⁽³⁾
11 - AI Disinformation	3,582 ⁽⁷⁾	3,405 ⁽⁶⁾	3,851 ⁽⁴⁾	3,377 ⁽¹⁰⁾	3,510 ⁽⁶⁾	3,512 ⁽⁶⁾	3,708 ⁽³⁾	3,483 ⁽⁸⁾
12 - Climate Disasters	3,511 ⁽⁸⁾	3,190 ⁽¹¹⁾	3,660 ⁽⁶⁾	3,472 ⁽⁷⁾	3,245 ⁽⁹⁾	3,512 ⁽⁶⁾	3,625 ⁽⁵⁾	3,517 ⁽⁷⁾

Table 6 - Short-term BRGE2025 average perceived GPR scores, measured on a 1-5 scale, by GPR group, firm size and Openness Score (OS), with superscript numbers indicating each group's ranking within the respective subdivision. Source: own elaboration.

¹⁷ To better analyze them, these results are presented separately for the short-term and long-term horizons.

Table 6 presents the short-term BRGE2025 results by GPR group, firm size and Openness Score (OS). The first column, *Short Term*, reports the average perceived relevance of each GPR group across the full sample of firms. The remaining columns disaggregate these results by company size, distinguishing between micro firms, SMEs and large firms and by degree of openness to foreign markets, measured through the Openness Score.

As was explained in the previous chapter, the values were calculated by averaging the survey responses collected in the BRGE2025. All values are expressed on a 1-5 scale, where higher values indicate a higher perceived relevance of that risk. These variations are also expressed by the different colors of each cell. For each column, lower values are shown in red, intermediate values in yellow or orange and higher values in green. The rankings are also expressed by the numbers in superscript, with (1) corresponding to the highest value of that indicator. This eases the interpretation of the table by showing the ranking of GPR groups within each firm profile, indicating risk relevance in each subdivision.

Focusing solely on the first column, overall short-term results, Financial Crisis emerges as the most relevant risk, with an average of 3,913, followed by Cyberattacks (3,875), Trade Wars (3,783) and Armed Conflict in Europe (3,696). These results suggest that Portuguese firms are particularly concerned with risks that may have direct and immediate effects on business activity. By contrast, Weaponized Migration (3,109), Nuclear, Biological and Chemical Threats (3,262) and Civil and Political Unrest (3,359) receive the lowest scores, indicating that firms perceive them as less immediate or less connected to their short-term operations.

Company Size The segmented results reveal relevant differences across firm profiles. These differences become clearer when comparing firms by size, as each group displays a distinct pattern of concerns:

- **Micro firms** tend to display relatively lower and more homogeneous scores across most categories, with a stronger emphasis on Financial Crisis (3,952). This suggests stronger sensitivity to financial instability, reflecting their higher vulnerability to liquidity constraints and lower capacity to absorb shocks;
- **SMEs** show a noticeable increase in concern with risks linked to external environments, such as Trade Wars (3,957), Cyberattacks (3,936) and AI Disinformation (3,851), suggesting greater exposure to digital and international dynamics;

- **Large firms** assign comparatively higher importance to Supply Chain Disruption (3,736) and maintain elevated concern across several categories, which is consistent with their deeper integration into global production and distribution networks.

Analyzed as a whole, the results suggest a gradual shift from more general and financially driven concerns among micro firms to risks more closely associated with operational complexity and international exposure among SMEs and large firms.

Level of Openness A similar progression can be observed when analyzing firms by Openness Score. As openness increases, firms tend to show growing concern with risks that are transmitted through international channels. Less open firms (OS = 0) display more moderate and balanced perceptions across categories, while firms with intermediate openness (OS = 1 and OS = 2) show increasing concern with Financial Crisis and Trade Wars. At the highest level of openness (OS = 3), concern becomes more pronounced for globally driven risks such as Global Armed Conflict (3,724). Overall, the results suggest that greater international exposure is associated with a broader and more intense perception of GPR.

4.1.2. Long Term

Having examined firms' short-term GPR perceptions, the analysis now turns to the long term horizon. The table below (**Table 7**) reports the long-term BRGE2025 results.

Group Of GPR	Long Term	Micro Firms	SMEs	Large Firms	OS = 0	OS = 1	OS = 2	OS = 3
1 - Global Armed Conflict	3,660 ⁽⁸⁾	3,513 ⁽⁸⁾	3,638 ⁽⁸⁾	3,681 ⁽⁵⁾	3,702 ⁽⁷⁾	3,513 ⁽⁹⁾	3,565 ⁽¹⁰⁾	3,640 ⁽⁶⁾
2 - Armed Conflict in Europe	3,732 ⁽⁶⁾	3,526 ⁽⁷⁾	3,702 ⁽⁷⁾	3,848 ⁽²⁾	3,733 ⁽⁵⁾	3,641 ⁽⁶⁾	3,696 ⁽⁶⁾	3,720 ⁽³⁾
3 - Nuclear, Biological and Chemical Threats	3,258 ⁽¹¹⁾	3,184 ⁽¹²⁾	3,065 ⁽¹²⁾	3,391 ⁽¹⁰⁾	3,159 ⁽¹²⁾	3,205 ⁽¹²⁾	2,957 ⁽¹¹⁾	3,480 ⁽⁸⁾
4 – Cyberattacks	3,953 ⁽²⁾	4,135 ⁽¹⁾	4,022 ⁽¹⁾	3,652 ⁽⁶⁾	4,047 ⁽¹⁾	4,000 ⁽²⁾	3,739 ⁽⁵⁾	3,720 ⁽³⁾
5 - Weaponized Migration	3,233 ⁽¹²⁾	3,342 ⁽¹¹⁾	3,261 ⁽¹¹⁾	3,109 ⁽¹²⁾	3,295 ⁽¹¹⁾	3,308 ⁽¹¹⁾	3,217 ⁽¹²⁾	3,000 ⁽¹²⁾
6 - Civil and Political Unrest	3,460 ⁽¹⁰⁾	3,421 ⁽¹⁰⁾	3,565 ⁽⁹⁾	3,348 ⁽¹¹⁾	3,477 ⁽¹⁰⁾	3,513 ⁽⁹⁾	3,609 ⁽⁸⁾	3,120 ⁽¹¹⁾
7 - Energy Shocks	3,587 ⁽⁹⁾	3,579 ⁽⁶⁾	3,565 ⁽⁹⁾	3,630 ⁽⁷⁾	3,500 ⁽⁹⁾	3,769 ⁽⁵⁾	3,652 ⁽⁷⁾	3,360 ⁽¹⁰⁾
8 - Financial Crisis	4,040 ⁽¹⁾	4,026 ⁽²⁾	4,000 ⁽³⁾	4,065 ⁽¹⁾	3,977 ⁽²⁾	4,103 ⁽¹⁾	4,043 ⁽²⁾	3,960 ⁽¹⁾
9 - Supply Chain Disruption	3,773 ⁽⁴⁾	3,711 ⁽⁵⁾	3,717 ⁽⁶⁾	3,717 ⁽⁴⁾	3,705 ⁽⁶⁾	3,846 ⁽³⁾	3,609 ⁽⁸⁾	3,600 ⁽⁷⁾
10 - Trade Wars	3,926 ⁽³⁾	3,816 ⁽³⁾	4,022 ⁽¹⁾	3,822 ⁽³⁾	3,841 ⁽³⁾	3,846 ⁽³⁾	4,136 ⁽¹⁾	3,800 ⁽²⁾
11 - AI Disinformation	3,733 ⁽⁵⁾	3,737 ⁽⁴⁾	3,935 ⁽⁴⁾	3,478 ⁽⁹⁾	3,818 ⁽⁴⁾	3,641 ⁽⁶⁾	3,870 ⁽³⁾	3,520 ⁽⁹⁾
12 - Climate Disasters	3,673 ⁽⁷⁾	3,447 ⁽⁹⁾	3,822 ⁽⁵⁾	3,614 ⁽⁸⁾	3,535 ⁽⁸⁾	3,632 ⁽⁸⁾	3,773 ⁽⁴⁾	3,680 ⁽⁵⁾

Table 7 - Long-term BRGE2025 average perceived GPR scores, measured on a 1-5 scale, by GPR group, firm size and Openness Score (OS), with superscript numbers indicating each group's ranking within the respective subdivision. Source: own elaboration.

Table 7 presents the long-term BRGE2025 results by GPR group, firm size and Openness Score (OS). This distinction is very relevant as some risks may be perceived as more immediate while others may be perceived as long term threats. As in the short-term analysis, values correspond to average survey scores on a 1-5 scale, where higher values indicate greater perceived relevance. The first column reports the full-sample average, while the remaining columns disaggregate the results by firm profile. The same color scale is used to rank GPR groups within each column, with red indicating lower values and green indicating higher values, creating a ranking of GPR groups within each firm profile, indicating risk relevance in each subdivision. These rankings are also expressed by the numbers in superscript, with (1) corresponding to the highest value of that indicator.

Focusing solely on the first column, overall long term results, Financial Crisis remains the most relevant risk, with an average score of 4,040, followed by Cyberattacks (3,953) and Trade Wars (3,926). These results suggest that, in the long term, Portuguese firms remain particularly concerned with risks that may affect financial stability, digital security and international trade. By contrast, Weaponized Migration (3,233), Nuclear, Biological and Chemical Threats (3,258) and Civil and Political Unrest (3,460) receive the lowest scores, indicating that firms perceive these risks as less relevant to their long-term outlook.

Compared with the short term results, the ranking remains relatively stable, which indicates continuity in firms' main concerns. Nevertheless, most long term values are slightly higher than their short term equivalents, suggesting that Portuguese firms tend to perceive GPR as more relevant when considering a longer time horizon.

Company Size This trend also extends to the variations across company sizes, i.e., although rankings remain stable, the vast majority of long term values are higher than their short term equivalents. This increase is particularly visible among micro firms, which report higher long-term scores in almost all GPR Groups, which suggests a broadening of their long-term concerns. For SMEs, the shift from short term to long term is more moderate, but still consistent, which indicates that these firms already perceive several external and internationally connected risks as relevant in the short term.

Large firms, however, show a different pattern. Their long-term scores do not increase as uniformly as those of micro firms or SMEs. Some risks become more relevant (particularly Energy Shocks, Weaponized Migration, Trade Wars, Armed Conflict in Europe and Financial Crisis), but other categories (such as Cyberattacks, Civil and Political Unrest and

Supply Chain Disruption), slightly decrease in the long term. This suggests that large firms may already internalize some operational risks in the short term, while their longer term concerns shift more clearly toward macroeconomic and energy-related risks.

Thus, these patterns suggest that firm size may affect the time horizon through which GPR is perceived. More importantly, it seems to indicate that, the larger the firms, the more internalized long term risks are in the short term, which may reflect more structured planning processes, greater exposure to complex operational environments and/or a stronger capacity to anticipate how geopolitical developments can affect business activity over time.

Level of Openness A similar progression can be observed when analyzing firms by Openness Score. Once again, rankings remain stable, but long term values are higher than their short term equivalents, a trend that is less intense in companies with higher Openness Score. While firms with OS = 0 show a general increase in long term perceived relevance across all GPR categories, firms with OS = 3 show greater stability between the short and long term. Therefore, the conclusion drafted for larger companies also applies to firms that are more open to foreign markets, i.e., the more open to foreign markets the firms, the more internalized long term risks are in the short term.

4.2. MW-GPR Index

This subsection presents the results of the MW-GPR Index and is organized in two steps. First, it presents the intermediate results of the index construction process, namely the construction and validation of the database and the classification of headline framing. Then, it presents the results of the final MW-GPR indicators, including Volume, Sentiment, Composite and Abnormal Sentiment. These, however, are not the final results of the study since they will still be combined with the results of the BRGE2025 in 4.3. *Combination of BRGE2025 and MW-GPR Index*.

This structure allows the final results to be interpreted in light of the methodological steps that produced them, while providing a clear transition from data construction to indicator interpretation.

4.2.1. Intermediate Results

The initial extraction produced 15.280 news articles, collected according to the instructions explained in the Methodology. The distribution of this initial sample (across media sources, GPR groups and keywords) can be seen in **Annex XIII**.

After the initial extraction phase was finished, the dataset was filtered in two separate stages, as explained in the previous chapter. The first consisted of using an AI model to remove the opinion pieces. To validate this procedure, a randomly selected test sample was manually reviewed and compared with the classifications produced by the AI model. The detailed validation results are presented in **Annex XIV**. The validation exercise showed a high level of agreement between the manual and automated classifications, which supported the use of this methodology for the full dataset.

After the removal of opinion pieces, the dataset was reduced from the previous 15.280 to 14.039 articles. The distribution of the dataset (across media sources, GPR groups and keywords) after the completion of this intermediate stage can be found in **Annex XV**.

The second filtering stage consisted of removing articles that were not relevant to the GPR category to which they had initially been assigned. This step was validated through manually reviewed test samples for each GPR group. These samples were then filled manually and compared with the classifications produced by the Python model. The detailed validation matrices can be found in **Annex XVI**. Overall, the validation results indicate that the model performed sufficiently well to justify its application to the complete dataset, although the filtering process should still be interpreted as a classification-based approximation rather than a perfect separation between relevant and irrelevant articles.

Following this second stage, the final dataset consisted of 2.262 news articles. This sample was the basis used for the construction of the MW-GPR Index. The distribution of the final dataset (across media sources, GPR groups and keywords) can be analyzed in **Table 8**.

The results show that the final dataset is unevenly distributed across GPR categories and media sources. This is relevant, as the MW-GPR Index partly depends on the visibility of each GPR category in the media sample and groups with a larger number of articles have a stronger presence in the volume component of the MW-GPR Index. Underrepresented categories, however, may still become important if they display stronger sentiment intensity.

Group of GPR	Key-word	Newspapers					Total	
		Expresso	Negócios	Público	Observador	Diário Notícias	by Keyword	by GPR Group
1 - Global Armed Conflict	1	185	8	20	32	45	290	521
	2	52	6	12	32	8	110	
	3	75	3	4	27	12	121	
2 - Armed Conflict in Europe	1	138	9	24	43	21	235	360
	2	57	0	0	0	2	59	
	3	49	4	5	3	5	66	
3 - Nuclear, Biological and Chemical Threats	1	19	0	1	2	1	23	44
	2	1	0	0	0	0	1	
	3	12	1	2	1	4	20	
4 - Cyberattacks	1	12	2	4	2	11	31	102
	2	3	11	12	1	0	27	
	3	14	3	1	10	16	44	
5 - Weaponized Migration	1	12	9	4	18	15	58	88
	2	5	3	4	4	6	22	
	3	1	0	3	2	2	8	
6 - Civil and Political Unrest	1	69	4	44	49	78	244	573
	2	61	12	44	55	95	267	
	3	31	1	10	7	13	62	
7 - Energy Shocks	1	5	3	3	5	1	17	23
	2	1	1	2	1	0	5	
	3	1	0	0	0	0	1	
8 - Financial Crisis	1	5	7	6	1	9	28	48
	2	4	7	2	2	5	20	
	3	0	0	0	0	0	0	
9 - Supply Chain Disruption	1	0	3	0	2	0	5	17
	2	0	6	1	0	2	9	
	3	0	2	1	0	0	3	
10 - Trade Wars	1	7	3	0	0	2	12	28
	2	9	3	0	0	0	12	
	3	2	1	0	0	1	4	
11 - AI Disinformation	1	0	0	0	3	0	3	38
	2	5	1	9	4	5	24	
	3	0	2	3	5	1	11	
12 - Climate Disasters	1	4	0	35	14	4	57	420
	2	31	2	61	18	30	142	
	3	43	32	84	32	30	221	
Total		913	149	401	375	424	2262	2262

Table 8 - Distribution and number of extracted articles by GPR group, keyword and Portuguese online newspaper. Source: own elaboration.

Framing Classification After the final sample was established, the articles were classified according to the framing (Concerning, Neutral or Dismissive) expressed in their headlines. The summary of the classification can be seen on the table below (**Table 9**).

Framing Classification	Number of Articles
Concerning	622
Neutral	1.630
Dismissive	10

Table 9 - Number of articles classified as Concerning, Neutral or Dismissive based on headline framing. Source: own elaboration.

The results show that most headlines in the final dataset were classified as **Neutral**, while a smaller but still substantial share was classified as **Concerning**. **Dismissive** framing, however, was almost absent. This suggests that Portuguese media coverage of the selected Groups of GPR was primarily descriptive rather than explicitly alarmist.

Sensitivity Analysis A sensitivity analysis was conducted to assess whether the coefficients used to calculate the **Sentiment** indicator had a significant impact on the final results. This analysis, presented in **Annex XVII**, tested the robustness of the selected coefficient structure. Since the results remained fairly stable across the tested alternatives, the chosen coefficients were accepted.

A similar sensitivity analysis was also conducted for the **Composite** indicator, in order to evaluate whether the weighting structure between Volume and Sentiment affected the final outcomes. The results of this analysis can be found in **Annex XVIII**. As the results were also fairly positive and did not suggest that the final indicator was excessively dependent on a specific coefficient choice, the selected coefficients were accepted.

4.2.2. MW-GPR Final Indicators' Results

Having finalized all the intermediary calculations, the values for the final indicators were computed and ranked. The table below (**Table 10**) presents the results of the indicators (Volume, Sentiment, Composite and Abnormal Sentiment) by GPR Group.

Group Of GPR	Volume	Sentiment	Composite	Abnormal Sentiment
1 - Global Armed Conflict	4,626 ⁽²⁾	3,135 ⁽⁵⁾	3,881 ⁽¹⁾	2,279 ⁽⁹⁾
2 - Armed Conflict in Europe	3,468 ⁽⁴⁾	3,165 ⁽⁴⁾	3,316 ⁽⁴⁾	2,414 ⁽⁸⁾
3 - Nuclear, Biological and Chemical Threats	1,194 ⁽⁸⁾	5,000 ⁽¹⁾	3,097 ⁽⁵⁾	3,549 ⁽⁵⁾
4 – Cyberattacks	1,612 ⁽⁵⁾	2,935 ⁽⁶⁾	2,273 ⁽⁷⁾	1,296 ⁽¹¹⁾
5 - Weaponized Migration	1,511 ⁽⁶⁾	2,691 ⁽⁹⁾	2,101 ⁽⁹⁾	3,707 ⁽⁴⁾
6 - Civil and Political Unrest	5,000 ⁽¹⁾	2,687 ⁽¹⁰⁾	3,844 ⁽²⁾	2,588 ⁽⁷⁾
7 - Energy Shocks	1,043 ⁽¹¹⁾	1,000 ⁽¹²⁾	1,022 ⁽¹²⁾	4,453 ⁽³⁾
8 - Financial Crisis	1,223 ⁽⁷⁾	3,205 ⁽³⁾	2,214 ⁽⁸⁾	4,462 ⁽²⁾
9 - Supply Chain Disruption	1,000 ⁽¹²⁾	2,828 ⁽⁷⁾	1,914 ⁽¹⁰⁾	5,000 ⁽¹⁾
10 - Trade Wars	1,079 ⁽¹⁰⁾	4,212 ⁽²⁾	2,646 ⁽⁶⁾	1,000 ⁽¹²⁾
11 - AI Disinformation	1,151 ⁽⁹⁾	2,169 ⁽¹¹⁾	1,660 ⁽¹¹⁾	3,059 ⁽⁶⁾
12 - Climate Disasters	3,899 ⁽³⁾	2,793 ⁽⁸⁾	3,346 ⁽³⁾	1,597 ⁽¹⁰⁾

Table 10 - MW-GPR Index results by GPR group and indicator, measured in a scale of 1-5, with superscript numbers indicating each group's ranking within the respective indicator. Source: own elaboration.

Table 10 reports the results of the four indicators of MW-GPR Index (Volume¹⁸, Sentiment¹⁹, Composite²⁰ and Abnormal Sentiment²¹) by GPR group. As explained in the methodology, all indicators were standardized on a 1-5 scale to allow comparison across GPR groups. Thus, higher values indicate greater relative media relevance within each indicator. The color coding follows the same logic of the two previous tables. For each column, lower values are shown in red, intermediate values in yellow or orange and higher values in green. Therefore, the colors should be interpreted vertically, as they show the ranking of GPR groups according to each one of the indicators. These rankings are also expressed by the numbers in superscript, with (1) corresponding to the highest value of that indicator.

¹⁸ **Volume** captures the relative number of articles published about each GPR group.

¹⁹ **Sentiment** measures the emotional tone expressed in the headlines.

²⁰ **Composite** combines Volume and Sentiment into a single indicator.

²¹ **Abnormal Sentiment** captures the unexpected variation in recent media relevance.

Volume Focusing on the Volume column, the results show that media coverage was concentrated in a small number of GPR groups. Civil and Political Unrest records the highest value, followed by Global Armed Conflict and Climate Disasters. This suggests that Portuguese online newspapers gave greater attention to risks associated with visible conflict, social instability and climate-related events. By contrast, Supply Chain Disruption, Energy Shocks and Trade Wars received far less coverage.

Sentiment Focusing on the Sentiment column, the results differ considerably from those observed for Volume. Nuclear, Biological and Chemical Threats records the highest sentiment score, despite having relatively low media coverage, followed by Trade Wars and Financial Crisis. This suggests that some risks were not necessarily covered very frequently, but when they appeared in the news, they were framed in a more concerning or risk-amplifying way. By contrast, Energy Shocks records the lowest sentiment score, followed by AI Disinformation and Civil and Political Unrest.

Composite Focusing on the Composite column, the results show a combination of the two previous indicators. Global Armed Conflict records the highest Composite score, closely followed by Civil and Political Unrest and Climate Disasters. By contrast, Energy Shocks, AI Disinformation and Supply Chain Disruption record the lowest Composite values. These results indicate that, despite some of these risks presenting relevant sentiment scores, they were not as prominent when media visibility and headline framing are considered jointly. Overall, the Composite indicator confirms that high media relevance depends on the interaction between frequency and tone, rather than on either dimension alone.

Abnormal Sentiment Focusing on the Abnormal Sentiment column, the results show that Supply Chain Disruption records the highest value, followed by Financial Crisis and Energy Shocks. This suggests that, although several of these risks received relatively low overall media coverage, they became more prominent than expected in the most recent period. By contrast, Trade Wars, Cyberattacks and Climate Disasters record the lowest Abnormal Sentiment values. This indicates that these risks did not experience a strong positive deviation from their expected recent media relevance.

4.3. Combination of BRGE2025 and MW-GPR Index (Final Results)

This subchapter presents the final results of the dissertation by bringing together the two main components of the analysis: the MW-GPR Index and the BRGE2025. Its purpose is to assess whether the GPR groups that were more visible, more strongly framed or more unexpectedly salient in Portuguese online media are also those perceived as more relevant by Portuguese firms. In doing so, this section addresses the central objective of the study: understanding the relationship between media coverage and corporate perceptions of GPR. The analysis is conducted separately for the short and long term results.

4.3.1. Short Term

The short-term results are analyzed first, as they allow the relationship between media coverage and firms' immediate perceptions of GPR to be assessed. The table below (**Table 11**) presents the short term correlation.

Summary Table	Short Term							
	Short Term	Micro Emp	PME	Grand Emp	OS = 0	OS = 1	OS = 2	OS = 3
Volume	-0,042	-0,049	-0,046	-0,018	-0,035	-0,198	-0,109	0,210
Sentiment	0,455	0,315	0,235	0,585	0,329	0,335	0,263	0,503
Composite	0,042	-0,084	-0,014	0,172	-0,011	-0,127	-0,039	0,371
Abnormal Sentiment	-0,294	-0,196	-0,518	-0,077	-0,354	-0,176	-0,210	-0,573

Table 11 - Short-term correlation coefficients (-1 to 1) between MW-GPR Index indicators and BRGE2025 perceived GPR scores, by firm size and Openness Score (OS). Source: own elaboration.

Table 11 presents the short term correlations between the MW-GPR Index indicators and the BRGE2025 perceived GPR scores. To clarify how the table should be interpreted, the top left value, -0,042, can be used as an example. This value corresponds to the Spearman Rank Correlation between the ranking of the Volume indicator from the MW-GPR Index, shown in **Table 10** and analyzed in **Section 4.2.2**, and the ranking of the Short Term dimension from the BRGE2025, shown in **Table 6** and analyzed in **Section 4.1.1**. Therefore, -0,042 reflects the degree of similarity between these two rankings. More specifically:

- The negative sign indicates that GPR groups with higher media coverage were less likely to be perceived as more relevant by firms;

- The value is very close to zero, which indicates that this relationship is weak, suggesting that Volume is not aligned with short term corporate GPR perception.

Values between -0,2 and 0,2 were discarded as irrelevant. This applies, not only to -0,042, but also to most Volume correlations, as well as to several Composite and Abnormal Sentiment results. These values should therefore be interpreted with caution, as they do not provide strong evidence of alignment between the corresponding MW-GPR indicator and firms' perceived relevance of GPR.

Volume Firstly, it seems relevant to analyze the results that stem from the Volume variable. Apart from an irrelevant positive value (OS = 3), all correlations are irrelevant (between -0,20 and 0,20) and most are very close to zero. This pattern is visible in the overall short-term result (-0,042), as well as among Micro firms (-0,049), SMEs (-0,046), large firms (-0,018) and firms with OS = 0 (-0,035). These results suggest that the Volume of articles published on a given GPR might not be aligned with the perceived relevance of that GPR among firms, which leads to the first conclusion of this study: **the amount of media coverage is not, on its own, a relevant predictor of firms' short-term GPR perceptions.**

This may suggest that, despite capturing media attention, Volume might ignore whether coverage is relevant for businesses or whether it affects them. Additionally, some risks may suffer from media normalization i.e., if a topic is constantly presented in the media, companies may become used to it. High coverage may, therefore, reflect background noise rather than increasing perceived urgency.

Sentiment The results that stem from Sentiment provide a much different analysis. All correlations are positive and relevant (between 0,20 and 1,00), with the strongest values being found for Large firms (0,585), firms with OS = 3 (0,503) and the overall short-term ranking (0,455). This indicates that the tone of media coverage is more closely aligned with corporate GPR perception than the sheer volume of news. These results suggest that the Sentiment expressed by articles published on a given GPR is aligned with perceived relevance of that GPR among firms, which leads to the second conclusion of the study: **the emotional tone expressed in the headlines might have influenced the perception of GPR of the Portuguese companies.**

This finding is particularly relevant, as Sentiment captures an interpretative dimension of media coverage, i.e., besides communicating the occurrence of a geopolitical event, the

headline of an article frames the event, suggesting uncertainty, gravity or potential consequences. It is also consistent with the theoretical framework developed by literature, particularly with the role of framing and emotional responses in the formation of risk perception.

As discussed previously, perceived risk is shaped not only by cognitive assessments of probability and impact, but also by emotional dimensions such as anxiety, dread and immediacy. Headlines with stronger sentiment may activate these emotional dimensions more directly, increasing the salience of certain risks in the minds of decision-makers.

Additionally, this result reinforces the idea that media influence on GPR perception operates through qualitative mechanisms. The sentiment expressed in headlines can affect how risks are mentally categorized, how severe they appear and how easily they are recalled when firms assess the geopolitical environment. This is especially important in the case of GPR, since these risks are often complex, distant and difficult to measure objectively. In such contexts, companies may rely more heavily on interpretative cues provided by the media, using headline tone as a shortcut to assess the importance and potential impact of each risk.

Composite As a weighted average of the two previous variables, the Composite indicator produces less clear results than Volume or Sentiment when analyzed separately, especially since it combines two dimensions of media coverage that behave differently. The results support this interpretation. The overall correlation is very weak (0,042) and most values remain between -0,2 and 0,2, indicating no relevant relationship. This leads to the third conclusion of the study: **the weaker results of the Composite do not necessarily mean that media coverage is unrelated to corporate GPR perception; rather, they confirm that the relationship depends on which aspect of media coverage is measured.**

Abnormal Sentiment Finally, some interesting conclusions can be drawn from the analysis of the results of the Abnormal Sentiment, as all of the values are below zero and most of them reveal a strong negative relationship. The strongest negative values appear for SMEs (-0,518) and firms with OS = 3 (-0,573), while the overall correlation is also significant (-0,294). This suggests that firms' short-term GPR perceptions might react inversely to Abnormal Sentiment, which leads to the fourth conclusion of this study: **sudden increases in media relevance are associated with lower, rather than higher, perceived relevance of GPR among firms.**

This result may indicate that companies do not react immediately to sudden changes in

the media treatment of GPR. While Abnormal Sentiment can capture temporary media intensification around a specific topic, firms may require more sustained exposure before incorporating that GPR into their strategic perception.

Firm Characteristics The analysis is also divided according to firm characteristics, particularly firm size and openness to foreign markets. Since Sentiment is the only MW-GPR Index indicator that presents meaningful correlations across all firm profiles, this part of the analysis focuses on this dimension.

Regarding **Company Size**, Large firms appear to show a stronger alignment with media sentiment than Micro companies and SMEs. This suggests that the relationship between Sentiment and corporate GPR perception becomes stronger as firm size increases, which leads to the fifth conclusion of this study: **the larger the company, the more its GPR perception tends to align with the emotional tone expressed in articles.**

This may indicate that larger firms are more responsive to the tone of geopolitical news, possibly because they tend to have more complex operations, broader market exposure and more structured mechanisms for monitoring external risks. In contrast, smaller firms may be more disconnected from GPR due to a stronger focus on more immediate and operational concerns, which can make their perceptions less aligned with broader media framing.

A similar pattern can be observed when considering the **Openness Score**. Companies with higher levels of international exposure show a stronger relationship between Sentiment and perceived GPR relevance, which leads to the sixth conclusion of this study: **the more internationally exposed a company is, the more its perception of GPR tends to align with the emotional tone expressed in Portuguese media headlines.**

This result is coherent with the nature of GPR itself, as firms involved in imports, exports or foreign direct investment are more likely to be affected by GPR and thus are bound to be more attentive to them.

These results reinforce the idea that media sentiment does not affect all companies in the same way. Instead, its relationship with corporate GPR perception appears to be conditioned by firm characteristics. This brings us to a broader conclusion: **the role of media sentiment in shaping GPR perception is not homogeneous across the Portuguese business landscape, with company size and international exposure appearing as relevant moderating factors in the relationship between emotional tone expressed by the media and perceived GPR.**

4.3.2. Long Term

After analyzing the short term correlations, the analysis now focus on examining whether the relationship between MW-GPR Index and BRGE2025 changes over a longer time horizon. The table below (**Table 12**) reports the long term correlation results.

Summary Table	Long Term							
	Long Term	Micro Emp	PME	Grand Emp	OS = 0	OS = 1	OS = 2	OS = 3
Volume	-0,259	-0,371	-0,144	-0,203	-0,147	-0,415	-0,179	-0,004
Sentiment	0,308	0,091	0,196	0,517	0,245	0,102	0,067	0,623
Composite	-0,196	-0,406	-0,116	0,007	-0,168	-0,422	-0,182	0,179
Abnormal Sentiment	-0,161	-0,119	-0,418	-0,056	-0,294	0,032	-0,329	-0,389

Table 12 - Long-term correlation coefficients (-1 to 1) between MW-GPR Index indicators and BRGE2025 perceived GPR scores, by firm size and Openness Score (OS). Source: own elaboration.

Table 12 presents the long term correlations between the MW-GPR Index indicators and the BRGE2025 perceived GPR scores. To clarify how the table should be interpreted, the top left value, -0,259, can be used as an example. This value corresponds to the Spearman Rank Correlation between the ranking of the Volume indicator from the MW-GPR Index, shown in **Table 10** and analyzed in **Section 4.2.2**, and the ranking of the Long Term dimension from the BRGE2025, shown in **Table 7** and analyzed in **Section 4.1.2**. Therefore, -0,259 reflects the degree of similarity between these two rankings.

Volume In the short term, the correlations are mostly close to zero. In the long term, however, the relationship becomes more negative. The overall correlation decreases to -0,259, while stronger negative values appear for Micro firms (-0,371) and firms with OS = 1 (-0,415). The previously positive correlation for firms with OS = 3 also disappears, falling from 0,210 in the short term to -0,004 in the long term. These results suggest that media volume is not aligned with firms' perception of GPR in either time horizon and that, in the long term, higher media visibility may even be associated with lower perceived relevance among some firm groups.

Sentiment In the short term, the relationship between the two variables was relevant across all types of companies. In the long term, Sentiment remains positive but becomes less consistent across firm profiles. The overall correlation decreases from 0,455 to 0,308, and the

values for micro firms, SMEs, OS = 1 and OS = 2 all decline. However, the relationship remains strong for large firms (0,517) and becomes even stronger for firms with OS = 3, increasing from 0,503 to 0,623. This suggests that the alignment between media sentiment and corporate GPR perception persists in the long term, mainly among larger and more internationally open firms, but is seriously weakened.

Composite Compared with the short term, long-term Composite correlations become weaker and more negative. This once again suggests that combining media visibility and media tone does not necessarily improve alignment with firms' perceptions.

Abnormal Sentiment In the short term, all correlations are negative. In the long term, most remain negative (SMEs [-0,418] and/or firms with OS = 2 [-0,329]), although the relationship becomes slightly less pronounced in several cases (overall correlation increases to -0,161) still show relevant negative correlations. This suggests that the relationship between abnormal sentiment and corporate GPR perception behaves similarly in short and long term.

Firm Characteristics Regarding firm characteristics, the long term results confirm one of the main patterns observed in the short term: the relationship between media sentiment and corporate GPR perception is stronger among larger and more internationally open firms. Large firms continue to show a strong correlation with Sentiment (0,517), while firms with OS = 3 present the highest long term value in the entire table (0,623). This suggests that firm exposure matters despite the time horizon.

Therefore, the results suggest that, overall, the relationship between the MW-GPR Index and the BRGE2025 does not differ substantially between the short and long term. The main patterns remain broadly similar across both time horizons, particularly the weak alignment of Volume, the stronger relevance of Sentiment, the mixed results of the Composite indicator and the negative behavior of Abnormal Sentiment. This leads to the last conclusion of this study: **the relationship between media coverage and corporate GPR perception appears to be relatively stable across time horizons, despite a slight tendency to weaker or more negative correlations in the longer term.**

5. Limitations and Future Paths of Study

Despite the contributions presented above, this study is not without limitations. The first concerns the process used to collect data to construct the MW-GPR Index. By relying on Google Search queries, the composition of the final database may have been biased, due to the platform's indexing criteria and search-ranking logic. As a result, it cannot be guaranteed that every relevant article published during the selected period was captured.

A second limitation is related to the scope of the media sample. By focusing solely on Portuguese online newspapers, other relevant sources of information (such as television, radio or social media) were excluded. Therefore, the MW-GPR Index shouldn't be interpreted as a complete representation of the full media environment, but as an indicator of Portuguese online press coverage.

Another important limitation relates to the classification of each headline according to the way it's framed. By relying solely on an LLM to classify the headline framing, the analysis introduced a degree of subjectivity and reduced reproducibility, that may have affected the results. Additionally, the decision to classify only headlines should also be acknowledged as a limitation, as headlines do not always reflect the tone, complexity or argument of the article. As a result, the Sentiment variable may not capture the complete sentiment of the articles.

The construction of the MW-GPR indicators also involved several assumptions, particularly regarding the coefficients attributed to date, outlet and framing. These coefficients were necessary to transform qualitative dimensions into measurable values, but they inevitably introduce subjectivity into the MW-GPR Index. Although the sensitivity analysis that was conducted proved that none of the coefficients dominated the Index and validated the values chosen, different coefficient structures could produce different results.

However, the most important limitation of this study lies in the fact that **the results are correlational, not causal**. Even when positive correlations are found, they cannot be interpreted as evidence that media directly caused higher corporate GPR perception, as it is possible that both media and firms were responding to the same external geopolitical events. Therefore, this study can demonstrate alignment between Portuguese media framing and corporate perceptions, but it cannot establish that the Portuguese media played a role in the perception of GPR by Portuguese firms, even if the former preceded the latter temporally.

These limitations, however, do not invalidate the findings. They solely define the boundaries within which the results should be interpreted. Moreover, they open the door for future research to refine and expand the analysis. As such, this dissertation shouldn't be understood

as a definitive causal explanation of how media shapes GPR perception, but as an exploratory and context-specific contribution that creates a foundation for further research.

Future studies could expand the media sample beyond Portuguese online newspapers and include television, radio, social media, international newspapers or specialized business intelligence sources. This would allow for a more complete understanding of the information environment through which firms interpret GPR. Additionally, future research could improve the measurement of sentiment by analyzing the full content of articles, rather than only headlines and by comparing different classification methods, such as LLM-based classification, dictionary-based approaches and supervised models.

Another relevant pathway would be to move beyond correlation by using longitudinal or event-based approaches, allowing researchers to better understand whether media sentiment precedes changes in corporate GPR perception or merely reflects broader geopolitical developments.

6. Conclusion

The main objective of this study was to analyze whether there is a relationship between Portuguese media coverage of GPR and the way Portuguese companies perceive those same GPR. Additionally, the dissertation aimed to understand whether corporate GPR perceptions are more closely associated with the amount of media coverage given to each GPR group, with the sentiment expressed in news headlines or with more complex indicators combining these dimensions. It also aimed to assess whether this relationship varies according to company size and level of openness to foreign markets.

To answer these questions, this dissertation constructed the MW-GPR Index, an LLM-assisted media-based indicator of GPR visibility and framing in Portuguese online newspapers, and compared it with the results of the BRGE2025, a GPR barometer focused on Portuguese companies and business owners. This comparison allowed the study to connect two dimensions that are often analyzed separately, media-based indicators of GPR and survey-based measures of corporate risk perception.

After analyzing the comparison between the results of the MW-GPR Index and the BRGE2025 results, it can be concluded that there is a connection between Portuguese media coverage and corporate perceptions of GPR. However, this relationship isn't as linear or straightforward as suggested by the literature. As explained before, existing studies indicate that the volume of coverage and emotional tone can influence the way GPR are perceived. This study, however, suggests that this connection depends on the specific media dimension considered.

The weaker correlation shown by Volume and Composite indicates that the sheer number of articles published about a certain GPR Group isn't necessarily sufficient to explain the way companies perceive GPR. By contrast, the stronger results obtained through Sentiment suggest that the tone and framing of the headlines may be connected to the GPR firms' perception. Headlines that express stronger concern, urgency or threat may make certain GPR more salient and easier to interpret as strategically relevant.

The Composite and Abnormal Sentiment indicators offered more limited results, since both combine Volume and Sentiment. This created a less clear interpretation, where the stronger relationship observed in Sentiment may have been partially diluted by the weaker relationship found in Volume. Abnormal Sentiment's weaker alignment with BRGE2025 suggests that firms may not immediately react to sudden changes in the media environment, especially when those changes are not sustained over time.

Another important finding is that firm characteristics also play an important role in the corporate perception of GPR. Larger companies and companies with higher levels of openness to foreign countries show a stronger alignment with media sentiment. This suggests that firms with more complex operations and greater international exposure may be more sensitive to the way GPR are framed in the media.

Taken together, these contributions position this dissertation as an original and relevant contribution to the study of GPR perception. By combining a tailored media-based index with firm-level survey evidence, the thesis offers a methodological approach that captures not only the visibility of GPR in the media, but also the role of framing in shaping corporate perceptions. Its focus on Portuguese firms and Portuguese online media further grounds the analysis in a specific national context, while producing findings with broader implications for the study of risk perception, media influence and corporate decision-making. For this reason, the dissertation stands as a meaningful contribution to a still underexplored field, showing that the corporate perception of GPR is influenced not only by the world as it is, but also by the way that world is narrated.

7. Appendix – Variables, Dimensions and Indicators Glossary

BRGE2025	MW-GPR Index
Short/Long Term: Measures the average short/long term perceived relevance of each GPR group among all surveyed firms, computed from BRGE2025 responses on a 1-5 scale. Micro Firms: Measures the average perceived relevance of each GPR group among micro firms (for short or long term), computed from their BRGE2025 responses on a 1-5 scale. SMEs: Measures the average perceived relevance of each GPR group among SMEs (for short or long term), computed from their BRGE2025 responses on a 1-5 scale. Large Firms: Measures the average perceived relevance of each GPR group among large firms (for short or long term), computed from their BRGE2025 responses on a 1-5 scale. OS = 0: Measures the average perceived relevance of each GPR group among firms with Openness Score equal to 0 (for short or long term), computed from their BRGE2025 responses on a 1-5 scale. OS = 1: Measures the average perceived relevance of each GPR group among firms with Openness Score equal to 1 (for short or long term), computed from their BRGE2025 responses on a 1-5 scale. OS = 2: Measures the average perceived relevance of each GPR group among firms with Openness Score equal to 2 (for short or long term), computed from their BRGE2025 responses on a 1-5 scale. OS = 3: Measures the average perceived relevance of each GPR group among firms with Openness Score equal to 3 (for short or long term), computed from their BRGE2025 responses on a 1-5 scale.	Volume: Measures the relative media visibility of each GPR group, computed from the number of relevant articles published and standardized (through a min-max normalization) on a 1-5 scale. Sentiment: Measures the relative tone and framing intensity of media coverage for each GPR group, computed from article-level sentiment scores based on date, outlet and framing coefficients and standardized (through a min-max normalization) on a 1-5 scale. $Sentiment_k = Date_k \times Outlet_k \times Framing_k$ • Date: Measures the recency component of article sentiment, computed by assigning a date coefficient to each article according to its publication period. • Outlet: Measures the outlet component of article sentiment, computed by assigning a coefficient according to the newspaper in which each article was published. • Framing: Measures the framing component of article sentiment, computed by assigning a coefficient according to the way each headline was classified. Composite: Measures the combined media relevance of each GPR group, computed by combining the standardized Volume and Sentiment indicators into a single score. Abnormal Sentiment: Measures the recent variation in media relevance for each GPR group, computed as the difference between observed and expected Composite values and standardized (through a min-max normalization) on a 1-5 scale.

Table 13 - Variables, Dimensions and Indicators Glossary. Source: own elaboration

8. References

- Allianz (2025). *Allianz Risk Barometer*. <https://commercial.allianz.com/content/dam/onemarketing/commercial/commercial/reports/Allianz-Risk-Barometer-2025.pdf>
- Althaus, S. L., Edy, J. A., & Phalen, P. F. (2001). Using Substitutes for Full-Text News Stories in Content Analysis: Which Text Is Best? *American Journal of Political Science*, 45(3), 707. <https://doi.org/10.2307/2669247>
- Andres-Escayola, E., Ghirelli, C., Molina, L., Pérez, J. J., & Vidal, E. (2022). Using newspapers for textual indicators: Which and how many? Banco de España. <https://ssrn.com/abstract=4252588>
- Bazerman, M. H., & Moore, D. A. (2013). *Judgment in Managerial Decision Making*. Wiley.
- Beazley (2025). *Spotlight on Geopolitical & Economic Uncertainty 2025*. <https://www.beazley.com/globalassets/reports/spotlight-on-geopolitical-and-economic-uncertainty-2025/geopolitical-and-economic-report-2025.pdf>
- Beck, U. (1992). *Risk Society: Towards a New Modernity*. Sage Publications.
- BlackRock (2025, December 15). *Geopolitical Risk Dashboard*. <https://www.blackrock.com/corporate/literature/whitepaper/geopolitical-risk-dashboard-december-2025.pdf>
- Bondarenko, Y., Lewis, V., Rottner, M., & Schöler, Y. S. (2024). *Geopolitical Risk Perceptions*. <https://doi.org/10.2139/ssrn.4955431>
- Caldara, D., & Iacoviello, M. (2022). Measuring Geopolitical Risk. *American Economic Review*, 112(4), 1194–1225. <https://doi.org/10.1257/aer.20191823>
- Cambridge Dictionary (2025a, December 24). *geopolitics*. @CambridgeWords. <https://dictionary.cambridge.org/us/dictionary/english-portuguese/geopolitics>
- Cambridge Dictionary (2025b, December 24). *risk*. @CambridgeWords. <https://dictionary.cambridge.org/us/dictionary/english-portuguese/risk>
- Carney, M. (2016, June 30). Uncertainty, the economy and policy [Review of Uncertainty, the economy and policy]. <https://www.bankofengland.co.uk/speech/2016/uncertainty-the-economy-and-policy>
- Cunha, I. F., & Peixinho, A. T. (2020). Análise dos Media. In *Imprensa da Universidade de Coimbra eBooks*. <https://doi.org/10.14195/978-989-26-1988-0>
- Gill, J., Tostevin, R., & Hussain, E. (2014, June). *Hazards and the Himalaya*. Academia.edu. https://www.academia.edu/31148522/Hazards_and_the_Himalaya

- Glowka, G., Hule, R., & Zehrer, A. (2024). Risk perception of SMEs: strategic risks, family-related risks, external risks. *Risk Management*, 26(4). <https://doi.org/10.1057/s41283-024-00148-2>
- Jones, W. J. (2025). *Influence of Media on Risk Perception During the COVID-19 Pandemic: Qualitative Case Study from Thailand*. *Public Health Policy and Laws Journal*, 11(3), 713–729. retrieved from https://so05.tci-thaijo.org/index.php/journal_law/article/view/280884 .
- Kapuściński, G., & Richards, B. (2016). News framing effects on destination risk perception. *Tourism Management*, 57, 234–244. <https://doi.org/10.1016/j.tourman.2016.06.017>
- Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., Kasperson, J. X., & Ratick, S. (1988). The Social Amplification of Risk: A Conceptual Framework. *Risk Analysis*, 8(2), 177–187. <https://doi.org/10.1111/j.1539-6924.1988.tb01168.x>
- KPMG (2025). *Top Geopolitical Risks 2025*. <https://assets.kpmg.com/content/dam/kpmg-sites/xx/pdf/2025/03/top-geopolitical-risks-2025-web.pdf>
- Lauriola, M., Di Cicco, G., & Savadori, L. (2024). Apocalypse now or later? Nuclear war risk perceptions mirroring media coverage and emotional tone shifts in Italian news. *Judgment and Decision Making*, 19. <https://doi.org/10.1017/jdm.2024.2>
- Menon, V. A., Kunnumpurath, B., & K, J. (2025). Media's Role In Shaping Perceptions Of Disaster Tourism: Analysing Viewer Responses To The Wayanad And Ankola Landslides. *Atna Journal of Tourism Studies*, 20(1), 191–212. <https://doi.org/10.12727/ajts.33.9>
- Moussaïd, M. (2013). Opinion Formation and the Collective Dynamics of Risk Perception. *PLoS ONE*, 8(12), e84592. <https://doi.org/10.1371/journal.pone.0084592>
- Niu, C., Jiang, Z., Liu, H., Yang, K., Song, X., & Li, Z. (2020). The influence of media consumption on public risk perception: a meta-analysis. *Journal of Risk Research*, 25(1), 1–27. <https://doi.org/10.1080/13669877.2020.1819385>
- Oliveira, F. M., Holanda, T. C., Silva, L. M. T., & Abreu, M. C. S. D. (2025). Media narrative on benefits and impacts of green hydrogen production: exploring environmental and social issues. *Cadernos EBAPE.BR*, 23(2). <https://doi.org/10.1590/1679-395120240043x>
- Pinto, K., Gonçalves, G., Sousa, C., & Sousa, A. (2024). Explorando a percepção de risco: o impacto de fatores individuais e a mediação do clima de segurança física em trabalhadores portugueses. *Psique*, XX(1). <https://doi.org/10.26619/2183-4806.xx.1.4>
- Porto Business School (2025). *Barómetro do Observatório do Risco Geopolítico para Empresas: Relatório Anual de Resultados 2025*. <https://143334250.fs1.hubspotusercontent-eu1.net/hubfs/143334250/Documentos%20IXH/Impact%20Research/Bar%C3%B3me->

tro%20Risco%20Geopol%C3%ADtico_V3.pdf

Rice, C. & Zegart, A.B. (2018). *Political Risk: How Businesses and Organizations Can Anticipate Global Insecurity*. New York, Ny: Twelve, Hachette Book Group.

Scholvin, S. (2016). *Geopolitics: An Overview of Concepts and Empirical Examples From International Relations*. <https://www.files.ethz.ch/isn/196701/wp91-Geopolitics.pdf>

Shen, L. S. (2025). How Firms' Perceptions of Geopolitical Risk Affect Investment. Federal Reserve Bank of Boston. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5137387

Singer, E., & Endreny, P. M. (1994). *Reporting on Risk: How the Mass Media Portray Accidents, Diseases, Other Hazards*. Russell Sage Foundation.

Slovic, P. (1987). Perception of Risk. *Science*, 236(4799), 280–285. <https://doi.org/10.1126/science.3563507>

Wahlberg, A. A. F., & Sjoberg, L. (2000). Risk perception and the media. *Journal of Risk Research*, 3(1), 31–50. <https://doi.org/10.1080/136698700376699>

World Economic Forum (2025). *The Global Risks Report 2025*. https://reports.weforum.org/docs/WEF_Global_Risks_Report_2025.pdf

Venade, J., & Grilo, F. (2023). A Matter of Minutes: Unexpected FOMC Communication and Fed Credibility. *The B.E. Journal of Macroeconomics*. <https://doi.org/10.1515/bejm-2024-0074>

Yu, M., & Wang, N. (2023). The Influence of Geopolitical Risk on International Direct Investment and Its Countermeasures. *Sustainability*, 15(3), 2522. <https://doi.org/10.3390/su15032522>

Annexes

Annex I: Methodology followed by the BRGE2025

“O Barómetro de Risco Geopolítico para Empresas foi concebido com o objetivo de avaliar a perceção das empresas relativamente ao risco geopolítico e ao seu impacto na atividade empresarial. A metodologia adotada assenta em três pilares principais:

1. Recolha de informação: O questionário foi estruturado de forma a captar, de maneira objetiva, as opiniões, preocupações e estratégias das empresas face aos atuais desenvolvimentos geopolíticos. As perguntas seguem um formato de identificação das situações que podem gerar risco mais elevado para o seu respetivo setor de atividade e, numa última fase, as ferramentas que melhor poderão mitigar e/ou resolver os riscos anteriormente identificados.
2. Representatividade: O inquérito dirige-se a empresas e outros atores económicos de diferentes dimensões, setores de atividade e perfil de internacionalização, de forma a obter uma visão abrangente e comparável entre realidades distintas. O convite à participação foi alargado a organizações nacionais e internacionais com presença em Portugal.
3. Tratamento e análise de dados: As respostas foram analisadas de forma agregada, garantindo a confidencialidade dos participantes. A análise combinou técnicas estatísticas descritivas com a interpretação qualitativa das respostas abertas, permitindo identificar padrões, riscos prioritários e eventuais medidas de mitigação adotadas pelas empresas.

Esta abordagem visa assegurar a fiabilidade dos resultados, refletindo de forma transparente as perceções do tecido empresarial sobre os riscos geopolíticos atuais e futuros. Para assegurar uma melhor qualidade na amostra, foi definida uma preocupação maior em relação à simplificação do inquérito, limitando ao máximo os riscos de erros de interpretação e diminuindo o tempo de resposta na sua totalidade.

O inquérito foi dividido em três áreas, cada uma delas correspondendo a uma pergunta. Para cada questão, o inquirido foi convidado a classificar riscos ou ferramentas, seguindo uma escala de Likert, em cinco níveis de relevância de “Muito Baixo” a “Muito Elevado” (para mais detalhes, ler o Anexo A, que detalha o texto do questionário):

- Classifique os riscos geopolíticos tendo em conta o possível impacto no seu setor de atividade, no curto prazo (1 ano) com a identificação do nível percecionado de risco para os 13 riscos apresentados (sendo o último a possibilidade de identificar “outros riscos”).
- Classifique os riscos geopolíticos tendo em conta o possível impacto no seu setor de atividade, no médio/longo prazo (a 3 anos) com a identificação do nível percecionado de risco para os 13 riscos apresentados (sendo o último a possibilidade de identificar “outros riscos”), considerando o médio/longo prazo. É de relevar que o espaço temporal considerado para esta questão foi limitado a três anos 33 pela elevada dinâmica dos acontecimentos geopolíticos globais, que condicionam uma análise a longo prazo, e pelo menor conhecimento geopolítico específico dos inquiridos, o que poderia afetar a qualidade das respostas.
- Classifique as ferramentas tendo em conta a sua possível eficácia na mitigação dos riscos geopolíticos importantes para o seu sector a 3 anos com a identificação do nível percecionado para as 10 ferramentas que poderão mitigar e/ou resolver os riscos anteriormente identificados (sendo a última a possibilidade de identificar “outros ferramentas”). Considerou-se o período de três anos, visto que um prazo inferior dificilmente permitiria apresentar resultados efetivos.

Na fase final, os inquiridos foram convidados a preencher alguns dados que permitiram fazer uma análise mais fina das respostas, considerando a sua dimensão, o seu setor e a sua exposição ao mercado internacional.

Foram recebidas 260 respostas de executivos de empresas baseadas em Portugal e com operação nacional e internacional, abrangendo uma análise setorial que inclui a indústria transformadora, empresas exportadoras e importadoras. Após tratamento, foram consideradas 185 respostas como válidas. O inquérito foi realizado entre 11 de janeiro e 8 de fevereiro deste ano, permitindo uma visão abrangente sobre a perceção do risco geopolítico e as estratégias de mitigação adotadas pelas empresas em Portugal.”

Annex II: Google Search Queries

GPR Group 1: Novos conflitos globais ou descontrolo dos atuais

- 1.X.1** - site:expresso.pt/ guerra AND militar AND internacional after:2024-02-08 before:2025-02-08
- 1.X.2** - site:expresso.pt/ armamento AND militar after:2024-02-08 before:2025-02-08
- 1.X.3** - site:expresso.pt/ NATO after:2024-02-08 before:2025-02-08
- 1.N.1** – site:jornaldenegocios.pt/ guerra AND militar AND internacional after:2024-02-08 before:2025-02-08
- 1.N.2** - site:jornaldenegocios.pt/ armamento AND militar after:2024-02-08 before:2025-02-08
- 1.N.3** - site:jornaldenegocios.pt/ NATO after:2024-02-08 before:2025-02-08
- 1.O.1** – site:observador.pt/ guerra AND militar AND internacional after:2024-02-08 before:2025-02-08
- 1.O.2** - site:observador.pt/ armamento AND militar after:2024-02-08 before:2025-02-08
- 1.O.3** - site:observador.pt/ NATO after:2024-02-08 before:2025-02-08
- 1.P.1** – site:publico.pt/ guerra AND militar AND internacional after:2024-02-08 before:2025-02-08
- 1.P.2** - site:publico.pt/ armamento AND militar after:2024-02-08 before:2025-02-08
- 1.P.3** - site:publico.pt/ NATO after:2024-02-08 before:2025-02-08
- 1.D.1** – site:dn.pt/ guerra AND militar AND internacional after:2024-02-08 before:2025-02-08
- 1.D.2** - site:dn.pt/ armamento AND militar after:2024-02-08 before:2025-02-08
- 1.D.3** - site:dn.pt/ NATO after:2024-02-08 before:2025-02-08

GPR Group 2: Novos conflitos na Europa ou descontrolo dos atuais

- 2.X.1** - site:expresso.pt/ guerra AND militar AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08
- 2.X.2** - site:expresso.pt/ armamento AND militar AND (Europa OR Ucrânia) after:2024-02-08 before:2025-02-08
- 2.X.3** - site:expresso.pt/ NATO AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08
- 2.N.1** – site:jornaldenegocios.pt/ guerra AND militar AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08

2.N.2 - site:jornaldenegocios.pt/ armamento AND militar AND (Europa OR Ucrânia) after:2024-02-08 before:2025-02-08

2.N.3 - site:jornaldenegocios.pt/ NATO AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08

2.O.1 – site:observador.pt/ guerra AND militar AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08

2.O.2 - site:observador.pt/ armamento AND militar AND (Europa OR Ucrânia) after:2024-02-08 before:2025-02-08

2.O.3 - site:observador.pt/ NATO AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08

2.P.1 – site:publico.pt/ guerra AND militar AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08

2.P.2 - site:publico.pt/ armamento AND militar AND (Europa OR Ucrânia) after:2024-02-08 before:2025-02-08

2.P.3 - site:publico.pt/ NATO AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08

2.D.1 – site:dn.pt/ guerra AND militar AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08

2.D.2 - site:dn.pt/ armamento AND militar AND (Europa OR Ucrânia) after:2024-02-08 before:2025-02-08

2.D.3 - site:dn.pt/ NATO AND (Ucrânia OR Europa) after:2024-02-08 before:2025-02-08

GPR Group 3: Explosão nuclear ou libertação de agentes biológicos ou químicos

3.X.1 - site:expresso.pt/ explosão AND nuclear after:2024-02-08 before:2025-02-08

3.X.2 - site:expresso.pt/ arma AND (biológico OR químico) after:2024-02-08 before:2025-02-08

3.X.3 - site:expresso.pt/ ameaça AND guerra AND nuclear after:2024-02-08 before:2025-02-08

3.N.1 – site:jornaldenegocios.pt/ explosão AND nuclear after:2024-02-08 before:2025-02-08

3.N.2 - site:jornaldenegocios.pt/ arma AND (biológico OR químico) after:2024-02-08 before:2025-02-08

3.N.3 - site:jornaldenegocios.pt/ ameaça AND guerra AND nuclear after:2024-02-08 before:2025-02-08

fore:2025-02-08

3.O.1 – site:observador.pt/ explosão AND nuclear after:2024-02-08 before:2025-02-08

3.O.2 - site:observador.pt/ arma AND (biológico OR químico) after:2024-02-08 before:2025-02-08

3.O.3 - site:observador.pt/ ameaça AND guerra AND nuclear after:2024-02-08 before:2025-02-08

3.P.1 – site:publico.pt/ explosão AND nuclear after:2024-02-08 before:2025-02-08

3.P.2 - site:publico.pt/ arma AND (biológico OR químico) after:2024-02-08 before:2025-02-08

3.P.3 - site:publico.pt/ ameaça AND guerra AND nuclear after:2024-02-08 before:2025-02-08

3.D.1 – site:dn.pt/ explosão AND nuclear after:2024-02-08 before:2025-02-08

3.D.2 - site:dn.pt/ arma AND (biológico OR químico) after:2024-02-08 before:2025-02-08

3.D.3 - site:dn.pt/ ameaça AND guerra AND nuclear after:2024-02-08 before:2025-02-08

GPR Group 4: Ciberataques de grande dimensão a infraestruturas críticas ou empresas

4.X.1 - site:expresso.pt/ ciberataque OR “ataque informático” after:2024-02-08 before:2025-02-08

4.X.2 - site:expresso.pt/ ataque AND digital AND (infraestrutura OR empresa) after:2024-02-08 before:2025-02-08

4.X.3 - site:expresso.pt/ cibersegurança after:2024-02-08 before:2025-02-08

4.N.1 – site:jornaldenegocios.pt/ ciberataque OR “ataque informático” after:2024-02-08 before:2025-02-08

4.N.2 - site:jornaldenegocios.pt/ ataque AND digital AND (infraestrutura OR empresa) after:2024-02-08 before:2025-02-08

4.N.3 - site:jornaldenegocios.pt/ cibersegurança after:2024-02-08 before:2025-02-08

4.O.1 – site:observador.pt/ ciberataque OR “ataque informático” after:2024-02-08 before:2025-02-08

4.O.2 - site:observador.pt/ ataque AND digital AND (infraestrutura OR empresa) after:2024-02-08 before:2025-02-08

4.O.3 - site:observador.pt/ cibersegurança after:2024-02-08 before:2025-02-08

4.P.1 – site:publico.pt/ ciberataque OR “ataque informático” after:2024-02-08 before:2025-02-08

4.P.2 - site:publico.pt/ ataque AND digital AND (infraestrutura OR empresa) after:2024-02-08 before:2025-02-08

4.P.3 - site:publico.pt/ cibersegurança after:2024-02-08 before:2025-02-08

4.D.1 – site:dn.pt/ ciberataque OR “ataque informático” after:2024-02-08 before:2025-02-08

4.D.2 - site:dn.pt/ ataque AND digital AND (infraestrutura OR empresa) after:2024-02-08 before:2025-02-08

4.D.3 - site:dn.pt/ cibersegurança after:2024-02-08 before:2025-02-08

GPR Group 5: Migrações em massa usadas como arma política

5.X.1 - site:expresso.pt/ migração AND segurança after:2024-02-08 before:2025-02-08

5.X.2 - site:expresso.pt/ fronteira AND migrante after:2024-02-08 before:2025-02-08

5.X.3 - site:expresso.pt/ refugiado AND crise after:2024-02-08 before:2025-02-08

5.N.1 – site:jornaldenegocios.pt/ migração AND segurança after:2024-02-08 before:2025-02-08

5.N.2 - site:jornaldenegocios.pt/ fronteira AND migrante after:2024-02-08 before:2025-02-08

5.N.3 - site:jornaldenegocios.pt/ refugiado AND crise after:2024-02-08 before:2025-02-08

5.O.1 – site:observador.pt/ migração AND segurança after:2024-02-08 before:2025-02-08

5.O.2 - site:observador.pt/ fronteira AND migrante after:2024-02-08 before:2025-02-08

5.O.3 - site:observador.pt/ refugiado AND crise after:2024-02-08 before:2025-02-08

5.P.1 – site:publico.pt/ migração AND segurança after:2024-02-08 before:2025-02-08

5.P.2 - site:publico.pt/ fronteira AND migrante after:2024-02-08 before:2025-02-08

5.P.3 - site:publico.pt/ refugiado AND crise after:2024-02-08 before:2025-02-08

5.D.1 – site:dn.pt/ migração AND segurança after:2024-02-08 before:2025-02-08

5.D.2 - site:dn.pt/ fronteira AND migrante after:2024-02-08 before:2025-02-08

5.D.3 - site:dn.pt/ refugiado AND crise after:2024-02-08 before:2025-02-08

GPR Group 6: Radicalização social e política, greves e motins civis violentos

6.X.1 - site:expresso.pt/ manifestação OR motim after:2024-02-08 before:2025-02-08

6.X.2 - site:expresso.pt/ greve after:2024-02-08 before:2025-02-08

6.X.3 - site:expresso.pt/ radicalização AND (social OR política) after:2024-02-08 before:2025-02-08

6.N.1 – site:jornaldenegocios.pt/ manifestação OR motim after:2024-02-08 before:2025-02-08

6.N.2 - site:jornaldenegocios.pt/ greve after:2024-02-08 before:2025-02-08

6.N.3 - site:jornaldenegocios.pt/ radicalização AND (social OR política) after:2024-02-08 before:2025-02-08

6.O.1 – site:observador.pt/ manifestação OR motim after:2024-02-08 before:2025-02-08

6.O.2 - site:observador.pt/ greve after:2024-02-08 before:2025-02-08

6.O.3 - site:observador.pt/ radicalização AND (social OR política) after:2024-02-08 before:2025-02-08

6.P.1 – site:publico.pt/ manifestação OR motim after:2024-02-08 before:2025-02-08

6.P.2 - site:publico.pt/ greve after:2024-02-08 before:2025-02-08

6.P.3 - site:publico.pt/ radicalização AND (social OR política) after:2024-02-08 before:2025-02-08

6.D.1 – site:dn.pt/ manifestação OR motim after:2024-02-08 before:2025-02-08

6.D.2 - site:dn.pt/ greve after:2024-02-08 before:2025-02-08

6.D.3 - site:dn.pt/ radicalização AND (social OR política) after:2024-02-08 before:2025-02-08

GPR Group 7: Choques energéticos ligados ao petróleo e gás natural

7.X.1 - site:expresso.pt/ abastecimento AND energia after:2024-02-08 before:2025-02-08

7.X.2 - site:expresso.pt/ crise AND energética after:2024-02-08 before:2025-02-08

7.X.3 - site:expresso.pt/ choque AND (petróleo OR gás) after:2024-02-08 before:2025-02-08

7.N.1 – site:jornaldenegocios.pt/ abastecimento AND energia after:2024-02-08 before:2025-02-08

7.N.2 - site:jornaldenegocios.pt/ crise AND energética after:2024-02-08 before:2025-02-08

7.N.3 - site:jornaldenegocios.pt/ choque AND (petróleo OR gás) after:2024-02-08 before:2025-02-08

7.O.1 – site:observador.pt/ abastecimento AND energia after:2024-02-08 before:2025-02-08

7.O.2 - site:observador.pt/ crise AND energética after:2024-02-08 before:2025-02-08

7.O.3 - site:observador.pt/ choque AND (petróleo OR gás) after:2024-02-08 before:2025-02-08

7.P.1 – site:publico.pt/ abastecimento AND energia after:2024-02-08 before:2025-02-08
7.P.2 - site:publico.pt/ crise AND energética after:2024-02-08 before:2025-02-08
7.P.3 - site:publico.pt/ choque AND (petróleo OR gás) after:2024-02-08 before:2025-02-08
7.D.1 – site:dn.pt/ abastecimento AND energia after:2024-02-08 before:2025-02-08
7.D.2 - site:dn.pt/ crise AND energética after:2024-02-08 before:2025-02-08
7.D.3 - site:dn.pt/ choque AND (petróleo OR gás) after:2024-02-08 before:2025-02-08

GPR Group 8: Crise financeira grave

8.X.1 - site:expresso.pt/ crise AND financeira after:2024-02-08 before:2025-02-08
8.X.2 - site:expresso.pt/ recessão after:2024-02-08 before:2025-02-08
8.X.3 - site:expresso.pt/ instabilidade AND financeira after:2024-02-08 before:2025-02-08
8.N.1 – site:jornaldenegocios.pt/ crise AND financeira after:2024-02-08 before:2025-02-08
8.N.2 - site:jornaldenegocios.pt/ recessão after:2024-02-08 before:2025-02-08
8.N.3 - site:jornaldenegocios.pt/ instabilidade AND financeira after:2024-02-08 before:2025-02-08
8.O.1 – site:observador.pt/ crise AND financeira after:2024-02-08 before:2025-02-08
8.O.2 - site:observador.pt/ recessão after:2024-02-08 before:2025-02-08
8.O.3 - site:observador.pt/ instabilidade AND financeira after:2024-02-08 before:2025-02-08
8.P.1 – site:publico.pt/ crise AND financeira after:2024-02-08 before:2025-02-08
8.P.2 - site:publico.pt/ recessão after:2024-02-08 before:2025-02-08
8.P.3 - site:publico.pt/ instabilidade AND financeira after:2024-02-08 before:2025-02-08
8.D.1 – site:dn.pt/ crise AND financeira after:2024-02-08 before:2025-02-08
8.D.2 - site:dn.pt/ recessão after:2024-02-08 before:2025-02-08
8.D.3 - site:dn.pt/ instabilidade AND financeira after:2024-02-08 before:2025-02-08

GPR Group 9: Disrupção das cadeias de abastecimento globais

9.X.1 - site:expresso.pt/ “cadeia de abastecimento” OR “cadeias de abastecimento” after:2024-02-08 before:2025-02-08
9.X.2 - site:expresso.pt/ crise AND (logística OR abastecimento) after:2024-02-08 before:2025-02-08
9.X.3 - site:expresso.pt/ "Canal de Suez" OR "Canal do Panamá" OR “Estreito de Ormuz” after:2024-02-08 before:2025-02-08

9.N.1 – site:jornaldenegocios.pt/ “cadeia de abastecimento” OR “cadeias de abastecimento”
after:2024-02-08 before:2025-02-08

9.N.2 - site:jornaldenegocios.pt/ crise AND (logística OR abastecimento) after:2024-02-08
before:2025-02-08

9.N.3 - site:jornaldenegocios.pt/ "Canal de Suez" OR "Canal do Panamá" OR “Estreito de Ormuz” after:2024-02-08 before:2025-02-08

9.O.1 – site:observador.pt/ “cadeia de abastecimento” OR “cadeias de abastecimento” after:2024-02-08 before:2025-02-08

9.O.2 - site:observador.pt/ crise AND (logística OR abastecimento) after:2024-02-08
before:2025-02-08

9.O.3 - site:observador.pt/ "Canal de Suez" OR "Canal do Panamá" OR “Estreito de Ormuz” after:2024-02-08 before:2025-02-08

9.P.1 – site:publico.pt/ “cadeia de abastecimento” OR “cadeias de abastecimento” after:2024-02-08 before:2025-02-08

9.P.2 - site:publico.pt/ crise AND (logística OR abastecimento) after:2024-02-08
before:2025-02-08

9.P.3 - site:publico.pt/ "Canal de Suez" OR "Canal do Panamá" OR “Estreito de Ormuz” after:2024-02-08 before:2025-02-08

9.D.1 – site:dn.pt/ “cadeia de abastecimento” OR “cadeias de abastecimento” after:2024-02-08 before:2025-02-08

9.D.2 - site:dn.pt/ crise AND (logística OR abastecimento) after:2024-02-08 before:2025-02-08

9.D.3 - site:dn.pt/ "Canal de Suez" OR "Canal do Panamá" OR “Estreito de Ormuz” after:2024-02-08 before:2025-02-08

GPR Group 10: Conflitos comerciais de elevada intensidade (EUA, China, UE)

10.X.1 - site:expresso.pt/ “guerra comercial” after:2024-02-08 before:2025-02-08

10.X.2 - site:expresso.pt/ tarifa AND internacional after:2024-02-08 before:2025-02-08

10.X.3 - site:expresso.pt/ conflito AND comércio after:2024-02-08 before:2025-02-08

10.N.1 – site:jornaldenegocios.pt/ “guerra comercial” after:2024-02-08 before:2025-02-08

10.N.2 - site:jornaldenegocios.pt/ tarifa AND internacional after:2024-02-08 before:2025-02-08

10.N.3 - site:jornaldenegocios.pt/ conflito AND comércio after:2024-02-08 before:2025-02-

08

- 10.O.1 – site:observador.pt/ “guerra comercial” after:2024-02-08 before:2025-02-08
- 10.O.2 - site:observador.pt/ tarifa AND internacional after:2024-02-08 before:2025-02-08
- 10.O.3 - site:observador.pt/ conflito AND comércio after:2024-02-08 before:2025-02-08
- 10.P.1 – site:publico.pt/ “guerra comercial” after:2024-02-08 before:2025-02-08
- 10.P.2 - site:publico.pt/ tarifa AND internacional after:2024-02-08 before:2025-02-08
- 10.P.3 - site:publico.pt/ conflito AND comércio after:2024-02-08 before:2025-02-08
- 10.D.1 – site:dn.pt/ “guerra comercial” after:2024-02-08 before:2025-02-08
- 10.D.2 - site:dn.pt/ tarifa AND internacional after:2024-02-08 before:2025-02-08
- 10.D.3 - site:dn.pt/ conflito AND comércio after:2024-02-08 before:2025-02-08

GPR Group 11: Desinformação gerada por Inteligência Artificial

- 11.X.1 - site:expresso.pt/ deepfake after:2024-02-08 before:2025-02-08
- 11.X.2 - site:expresso.pt/ desinformação AND (“Inteligência Artificial” OR AI OR IA) after:2024-02-08 before:2025-02-08
- 11.X.3 - site:expresso.pt/ manipulação AND informação after:2024-02-08 before:2025-02-08
- 11.N.1 – site:jornaldenegocios.pt/ deepfake after:2024-02-08 before:2025-02-08
- 11.N.2 - site:jornaldenegocios.pt/ desinformação AND (“Inteligência Artificial” OR AI OR IA) after:2024-02-08 before:2025-02-08
- 11.N.3 - site:jornaldenegocios.pt/ manipulação AND informação after:2024-02-08 before:2025-02-08
- 11.O.1 – site:observador.pt/ deepfake after:2024-02-08 before:2025-02-08
- 11.O.2 - site:observador.pt/ desinformação AND (“Inteligência Artificial” OR AI OR IA) after:2024-02-08 before:2025-02-08
- 11.O.3 - site:observador.pt/ manipulação AND informação after:2024-02-08 before:2025-02-08
- 11.P.1 – site:publico.pt/ deepfake after:2024-02-08 before:2025-02-08
- 11.P.2 - site:publico.pt/ desinformação AND (“Inteligência Artificial” OR AI OR IA) after:2024-02-08 before:2025-02-08
- 11.P.3 - site:publico.pt/ manipulação AND informação after:2024-02-08 before:2025-02-08
- 11.D.1 – site:dn.pt/ deepfake after:2024-02-08 before:2025-02-08
- 11.D.2 - site:dn.pt/ desinformação AND (“Inteligência Artificial” OR AI OR IA) after:2024-

02-08 before:2025-02-08

11.D.3 - site:dn.pt/ manipulação AND informação after:2024-02-08 before:2025-02-08

GPR Group 12: Catástrofes climáticas extremas

12.X.1 - site:expresso.pt/ clima AND (catástrofe OR extremo) after:2024-02-08 before:2025-02-08

12.X.2 - site:expresso.pt/ “incêndio florestal” OR “incêndios florestais” after:2024-02-08 before:2025-02-08

12.X.3 - site:expresso.pt/ “alteração climática” OR “alterações climáticas” after:2024-02-08 before:2025-02-08

12.N.1 – site:jornaldenegocios.pt/ clima AND (catástrofe OR extremo) after:2024-02-08 before:2025-02-08

12.N.2 - site:jornaldenegocios.pt/ “incêndio florestal” OR “incêndios florestais” after:2024-02-08 before:2025-02-08

12.N.3 - site:jornaldenegocios.pt/ “alteração climática” OR “alterações climáticas” after:2024-02-08 before:2025-02-08

12.O.1 – site:observador.pt/ clima AND (catástrofe OR extremo) after:2024-02-08 before:2025-02-08

12.O.2 - site:observador.pt/ “incêndio florestal” OR “incêndios florestais” after:2024-02-08 before:2025-02-08

12.O.3 - site:observador.pt/ “alteração climática” OR “alterações climáticas” after:2024-02-08 before:2025-02-08

12.P.1 – site:publico.pt/ clima AND (catástrofe OR extremo) after:2024-02-08 before:2025-02-08

12.P.2 - site:publico.pt/ “incêndio florestal” OR “incêndios florestais” after:2024-02-08 before:2025-02-08

12.P.3 - site:publico.pt/ “alteração climática” OR “alterações climáticas” after:2024-02-08 before:2025-02-08

12.D.1 – site:dn.pt/ clima AND (catástrofe OR extremo) after:2024-02-08 before:2025-02-08

12.D.2 - site:dn.pt/ “incêndio florestal” OR “incêndios florestais” after:2024-02-08 before:2025-02-08

12.D.3 - site:dn.pt/ “alteração climática” OR “alterações climáticas” after:2024-02-08

before:2025-02-08

Annex III: Extraction AI Prompt

You are a meticulous data extraction assistant. Your task is to parse an input HTML document (Google SERP–style page, including inline/visible dates near each result) and produce a clean Excel file with two columns, **in this exact order**:

- **publish_date** (formatted as DD/MM/YYYY)
- **link** (the canonical URL for each organic search result)

You must:

1. Extract all organic result URLs (skip ads, “People also ask”, carousels, filters, account widgets, navigation chrome, etc.).
2. For each result, identify the human visible date text that pertains to the article (e.g., 24/02/2024, 06/10/2024, 04/08/2024). Commonly appears in the snippet area near the site name; ignore dates in script/style tags and telemetry.
3. Normalize the publish date to **DD/MM/YYYY**. If the day/month order is ambiguous, assume DD/MM/YYYY (Portuguese locale) unless the token clearly matches YYYY-MM-DD or “D MMM YYYY” / “MMM D, YYYY” formats.
4. Deduplicate exact duplicate links; keep the first valid date encountered.
5. Output an Excel (.xlsx) file named **Results.xlsx**, with a single sheet named **results** and columns exactly **publish_date, link** (publish_date must be the first column).
6. If a date for a given link is not confidently found, leave publish_date blank (don’t invent).
7. Validate that the file is created and non-empty. Report how many rows were written.
8. If the input is a .docx that contains raw HTML text, read the full text and treat it as HTML source.
9. Be resilient to truncated class names and nested spans; rely on semantic cues (anchor tags for results, nearby date patterns).
10. Do not include non-result URLs (tracking pixels, gen_204, pagead, accounts.google.com, navigation anchors, share links, etc.).

User Instructions (place in “user”):

INPUT: I will supply a file named 12.D.2.docx containing the full HTML of a Google search results page.

TASK: Extract every organic result's publish date and URL, then export to Results.xlsx with **publish_date as the first column**, formatted as DD/MM/YYYY.

DELIVERABLES:

- The file Results.xlsx with columns, in this exact order: **publish_date, link** (DD/MM/YYYY).
- A short summary: number of links found, number of rows written, and any links without dates.

Tool Use / Implementation Guidance:

1. Load the input

- If it's a .docx that contains the HTML, extract full text (no formatting) and store it as a string html.

2. Parse HTML

- Use a robust HTML parser (e.g., BeautifulSoup/lxml).
- Identify organic result anchors. Fallback heuristics:
 - `<a>` inside containers with result-like structures; titles usually in `<h3 class="LC20lb">` within a clickable `<a>`.
 - Exclude links whose href contains: accounts.google.com, pagead, googleusercontent, gen_204, /search?, webhp, /preferences, /https-ervice/, support.google, gstatic, lens.google, maps.google.

3. Clean URLs

- For links routed via `/url?` patterns, extract the real target from the `q` or `url` parameter and URL-decode; otherwise keep the href.

4. Find dates near each result

- From each result container, look for sibling/descendant text nodes matching date regex patterns typical for PT/Europe:
 - `\b\d{1,2}/\d{1,2}/\d{4}\b`

- `\b\d{1,2}-\d{1,2}-\d{4}\b`
- `\b\d{4}-\d{1,2}-\d{1,2}\b`
- `\b\d{1,2}\s+(jan|fev|mar|abr|mai|jun|jul|ago|set|out|nov|dez)\w*\s+\d{4}\b` (case-insensitive)
- Prefer the closest date text to the result anchor/snippet; ignore unrelated dates.

5. **Normalize dates**

- For DD/MM/YYYY or DD-MM-YYYY, parse as day-first and output DD/MM/YYYY.
- For YYYY-MM-DD, convert to DD/MM/YYYY.
- For PT month-name formats, map jan → 01 ... dez → 12 and output DD/MM/YYYY.

6. **Deduplicate**

- Use the canonical link as key. Keep the first encountered non-empty date.

7. **Export Excel**

- Create Results.xlsx with sheet results.
- **Write columns in this exact order: publish_date, link.**
- Ensure UTF-8 handling.

8. **Report**

- Print: total results parsed, unique links, rows written, count of rows with blank publish_date.

Annex IV: Extracted HTML

<https://drive.google.com/drive/folders/1BHW6xFs4EtuRzgsVvTPCXuGrD3KSbmeQ?usp=sharing>

Annex V: Opinion Piece Removal AI Prompt

You are a meticulous research assistant specialized in media content analysis for academic research in economics and political economy.

User

Prompt:

You will receive an Excel file in which each row corresponds to one news article from an online newspaper.

Your task is to classify each article strictly into one of the following two categories:

- **Opinion Piece**
- **Not Opinion Piece**

Definition of categories:

An article must be classified as **Opinion Piece** if at least one of the following conditions holds:

- It is explicitly labeled as opinion, commentary, column, editorial, analysis, or signed commentary.
- The article primarily expresses personal judgments, normative statements, interpretations, or viewpoints of an author.
- The tone is argumentative or evaluative rather than descriptive.
- The article is written in the first person or clearly reflects the author's subjective stance.
- The article aims to persuade, interpret, or assess events rather than report them.

An article must be classified as **Not Opinion Piece** if it:

- Is primarily factual, descriptive, or informative.
- Reports events, data, statements, or developments without explicit judgment or personal interpretation.
- Corresponds to news reporting, briefings, updates, or explanatory journalism.
- Focuses on “what happened” rather than “what it means” or “what should be done.”



Important

rule:

If you are unsure, default to **Not Opinion Piece**. Do not overclassify opinion.

Instructions:

1. Read each article independently.
 2. Base your classification on:
 - the title,
 - the article text or excerpt,
 - and any metadata provided (e.g., section name, author).
 3. Do not infer opinion based on topic alone — only on writing style and intent.
 4. Do not invent information.
 5. Do not change, remove, or reorder any rows.
-

Output

format:

Return an Excel file with **two columns**:

- **Column A:** Link of the article
- **Column B:** Classification (values must be exactly one of the following: **Opinion Piece** or **Not Opinion Piece**)

Do **not** add explanations, comments, or summaries in the Excel file.

After producing the Excel file, provide a short summary in plain text stating:

- total number of articles processed,
- number classified as **Opinion Piece**,
- number classified as **Not Opinion Piece**.

Annex VI: Opinion Piece Removal Stata Validation Sample Prompt

/*****

STRATIFIED VALIDATION SAMPLE

Strata order:

- 1) Predicted Class
- 2) GPR Category
- 3) Paper
- 4) Custom quarter bucket

Output: validation_sample.xlsx

*****/

clear all

set more off

version 16.0

/*****

0) USER SETTINGS

*****/

local infile "C:\Users\Pedro Carvalho\Desktop\Stata Opinion Piece.xlsx"

local sheet "Sheet1"

local linkvar "LINK"

local classvar "CLASS"

local gprvar "GROUP"

local papervar "PAPER"

local datevar "DATE"

local n_opinion 300

local n_notopinion 300

local seed 20260305

```
/*****
```

```
1) IMPORT
```

```
*****/
```

```
import excel using "`infile'", sheet("`sheet'") firstrow clear
```

```
keep `linkvar' `classvar' `gprvar' `papervar' `datevar'
```

```
/*****
```

```
2) CLEAN CLASS + BINARY INDICATOR
```

```
*****/
```

```
capture confirm string variable `classvar'
```

```
if !_rc {
```

```
    replace `classvar' = trim(`classvar')
```

```
}
```

```
gen byte class_bin = .
```

```
replace class_bin = 1 if `classvar' == "Opinion Piece"
```

```
replace class_bin = 0 if `classvar' == "Not Opinion Piece"
```

```
assert class_bin != .
```

```
label define classlbl 0 "Not Opinion Piece" 1 "Opinion Piece"
```

```
label values class_bin classlbl
```

```
/*****
```

```
3) CLEAN PAPER + GPR
```

```
*****/
```

```
capture confirm string variable `papervar'
```

```
if !_rc {
```

```
    replace `papervar' = trim(`papervar')
```

```
}
```

```
capture confirm string variable `gprvar'
```

```

if !_rc {
    replace `gprvar' = trim(`gprvar')
}

/* Keep only the 5 papers you specified */
keep if inlist(`papervar', "Expresso", "Público", "Negócios", "Observador", "Diário Notí-
cias")

/* Keep only Group 1..Group 12 */
capture confirm string variable `gprvar'
if !_rc {
    gen byte gpr_ok = regexm(`gprvar', "^Group ([1-9] | 1[0-2])$")
}
else {
    gen byte gpr_ok = inrange(`gprvar', 1, 12)
}
keep if gpr_ok==1
drop gpr_ok

/*****
4) PARSE DATE AND CREATE CUSTOM QUARTER BUCKETS
*****/
gen double date_d = .

capture confirm numeric variable `datevar'
if _rc == 0 {
    replace date_d = `datevar'
}
else {
    replace date_d = daily(`datevar', "DMY")
}

format date_d %td

```

```

drop if missing(date_d)

/* Define your custom quarter boundaries */
local d_q1_start = daily("08/02/2024","DMY")
local d_q1_end   = daily("07/05/2024","DMY")

local d_q2_start = daily("08/05/2024","DMY")
local d_q2_end   = daily("07/08/2024","DMY")

local d_q3_start = daily("08/08/2024","DMY")
local d_q3_end   = daily("07/11/2024","DMY")

local d_q4_start = daily("08/11/2024","DMY")
local d_q4_end   = daily("08/02/2025","DMY")

/* Create quarter bucket variable */
gen byte qbucket = .
replace qbucket = 1 if inrange(date_d, `d_q1_start', `d_q1_end')
replace qbucket = 2 if inrange(date_d, `d_q2_start', `d_q2_end')
replace qbucket = 3 if inrange(date_d, `d_q3_start', `d_q3_end')
replace qbucket = 4 if inrange(date_d, `d_q4_start', `d_q4_end')

drop if missing(qbucket)

label define qlbl 1 "08/02/2024–07/05/2024" ///
                2 "08/05/2024–07/08/2024" ///
                3 "08/08/2024–07/11/2024" ///
                4 "08/11/2024–08/02/2025"

label values qbucket qlbl

/*****
4.5) CHECK WHETHER 300 PER CLASS IS FEASIBLE

```

```

*****/

tab class_bin

count if class_bin==1
local available_opinion = r(N)

count if class_bin==0
local available_notopinion = r(N)

if `available_opinion' < `n_opinion' {
    di as error "WARNING: Fewer than `n_opinion' Opinion Piece articles available after
filtering."
    di as error "Available Opinion Piece articles: `available_opinion'"
}

if `available_notopinion' < `n_notopinion' {
    di as error "WARNING: Fewer than `n_notopinion' Not Opinion Piece articles available
after filtering."
    di as error "Available Not Opinion Piece articles: `available_notopinion'"
}

/*****
5) STRATIFIED RANDOM SAMPLING
*****/

set seed `seed'
gen double u = runiform()

bysort class_bin `gprvar' `papervar' qbucket: gen long Nh = _N
bysort class_bin: egen long Nclass = total(Nh)

gen long n_target = .
replace n_target = `n_opinion' if class_bin==1
replace n_target = `n_notopinion' if class_bin==0

```

```

gen long nh = round(n_target * Nh / Nclass)

/* Minimum 1 per non-empty stratum to encourage heterogeneity */
replace nh = 1 if nh==0 & Nh>0 & n_target>=100

bysort class_bin `gprvar' `papervar' qbucket (u): gen long r = _n
gen byte selected = (r <= nh)

/*****
6) TOP-UP / TRIM TO EXACT N PER CLASS
*****/

/* Count how many selected in each class */
bysort class_bin: egen long n_selected = total(selected)

/* Class-specific deficits */
gen long deficit = .
replace deficit = `n_opinion' - n_selected if class_bin==1
replace deficit = `n_notopinion' - n_selected if class_bin==0
replace deficit = 0 if deficit < 0

/* Top-up from remaining unsampled observations */
gen double u_topup = runiform() if selected==0
bysort class_bin (u_topup): gen long r_topup = _n if selected==0
replace selected = 1 if selected==0 & r_topup <= deficit

/* Trim if overshoot */
gen double u_trim = runiform() if selected==1
bysort class_bin (u_trim): gen long r_trim = _n if selected==1

replace selected = 0 if class_bin==1 & selected==1 & r_trim > `n_opinion'
replace selected = 0 if class_bin==0 & selected==1 & r_trim > `n_notopinion'

```

```

/*****
7) EXPORT SAMPLE FOR MANUAL CODING
*****/

preserve
keep if selected==1

gen str20 Predicted_Class = cond(class_bin==1, "Opinion Piece", "Not Opinion Piece")
gen str20 Human_Class = ""

order `linkvar' Predicted_Class Human_Class `gprvar' `papervar' qbucket date_d

export excel using "validation_sample.xlsx", firstrow(variables) replace

restore

display "Done. Exported: validation_sample.xlsx"
display "Seed used: `seed'"

/* Final check */
count if selected==1 & class_bin==1
di as text "Final Opinion Piece sample size: " r(N)

count if selected==1 & class_bin==0
di as text "Final Not Opinion Piece sample size: " r(N)

```

Annex VII: Opinion Piece Removal Validation Sample Statistical Details

AI Classification	Testing Sample
Not Opinion Piece	300
Opinion Piece	300

Annex VIII: Relevancy Classification Python Prompt

```
!pip install pandas scikit-learn sentence-transformers openpyxl
```

```
[ENTER]
```

```
import os
```

```
os.listdir()
```

```
[ENTER]
```

```
import pandas as pd
```

```
import numpy as np
```

```
from sentence_transformers import SentenceTransformer
```

```
from sklearn.model_selection import train_test_split
```

```
from sklearn.linear_model import LogisticRegression
```

```
from sklearn.metrics import classification_report, confusion_matrix, precision_recall_curve
```

```
from sklearn.utils.class_weight import compute_class_weight
```

```
#
```

```
=====
```

```
=====
```

```
# 1. LOAD TRAINING SAMPLE
```

```
#
```

```
=====
```

```
=====
```

```
training_sample = pd.read_excel("Training Sample.xlsx")
```

```
#
```

```
=====
```

```
=====
```

```
# 2. CLEAN DATA
```

```
#
```

```
=====
```

```
=====
```

```

training_sample["headline"] = training_sample["headline"].fillna("").astype(str).str.strip()
training_sample["label"] = training_sample["label"].fillna("").astype(str).str.strip()

# Keep only valid rows
training_sample = training_sample[training_sample["headline"].str.len() > 0].copy()
training_sample = training_sample[training_sample["label"].isin(["Relevant", "Irrelevant"])].copy()

# Convert labels to binary
training_sample["label_binary"] = training_sample["label"].map({
    "Relevant": 1,
    "Irrelevant": 0
})

# Model input
training_sample["model_text"] = training_sample["headline"]

print("Training sample size:", len(training_sample))
print(training_sample["label"].value_counts())

#
=====
====

# 3. TRAIN / VALIDATION SPLIT
#
=====
====

train_data, validation_data = train_test_split(
    training_sample,
    test_size=0.2,
    random_state=42,
    stratify=training_sample["label_binary"]
)

```

```

#
=====

====

# 4. CREATE EMBEDDINGS
#
=====

====

embedding_model = SentenceTransformer("all-MiniLM-L6-v2")

X_train = embedding_model.encode(train_data["model_text"].tolist(), show_pro-
gress_bar=True)
X_validation = embedding_model.encode(validation_data["model_text"].tolist(),
show_progress_bar=True)

y_train = train_data["label_binary"].values
y_validation = validation_data["label_binary"].values

#
=====

====

# 5. HANDLE CLASS IMBALANCE
#
=====

====

classes = np.unique(y_train)
weights = compute_class_weight(class_weight="balanced", classes=classes, y=y_train)
class_weight_dict = {cls: w for cls, w in zip(classes, weights)}

#
=====

====

# 6. TRAIN CLASSIFIER

```

```

#
=====

====
classifier = LogisticRegression(
    max_iter=3000,
    class_weight=class_weight_dict,
    random_state=42
)

classifier.fit(X_train, y_train)

#
=====

====
# 7. VALIDATE
#
=====

====
validation_probabilities = classifier.predict_proba(X_validation)[:, 1]
validation_predictions_050 = (validation_probabilities >= 0.5).astype(int)

print("\n=== RESULTS WITH THRESHOLD = 0.50 ===")
print(confusion_matrix(y_validation, validation_predictions_050))
print(classification_report(y_validation, validation_predictions_050, digits=3, zero_division=0))

#
=====

====
# 8. CHOOSE A BETTER THRESHOLD
#
=====

====

```

```

best_threshold = 0.5
print(f"\nChosen threshold: {best_threshold:.3f}")

print(f"\nChosen threshold: {best_threshold:.3f}")

validation_predictions_tuned = (validation_probabilities >= best_threshold).astype(int)

print("\n=== RESULTS WITH TUNED THRESHOLD ===")
print(confusion_matrix(y_validation, validation_predictions_tuned))
print(classification_report(y_validation, validation_predictions_tuned, digits=3, zero_division=0))

#
=====

===
# 9. SAVE VALIDATION OUTPUT
#
=====

===
validation_output = validation_data.copy()
validation_output["predicted_probability"] = validation_probabilities
validation_output["predicted_label_050"] = np.where(validation_predictions_050 == 1,
"Relevant", "Irrelevant")
validation_output["predicted_label_tuned"] = np.where(validation_predictions_tuned == 1,
"Relevant", "Irrelevant")

validation_output.to_csv("validation_results_first_gpr_type.csv", index=False)
print("\nSaved: validation_results_first_gpr_type.csv")

```

```

#
=====

====

# CLASSIFY THE FULL SAMPLE
#
=====

====

import pandas as pd
import numpy as np

# 1. Load full sample
full_sample = pd.read_excel("Full Sample.xlsx")

# 2. Clean text
full_sample["clean_headline"] = full_sample["clean_head-
line"].fillna("").astype(str).str.strip()
full_sample = full_sample[full_sample["clean_headline"].str.len() > 0].copy()

# 3. Use the same text field structure as in training
full_sample["model_text"] = full_sample["clean_headline"]

print("Rows to classify:", len(full_sample))

# 4. Create embeddings
X_full = embedding_model.encode(
    full_sample["model_text"].tolist(),
    show_progress_bar=True
)

# 5. Predict with threshold = 0.50
threshold = 0.50
full_probabilities = classifier.predict_proba(X_full)[:, 1]
full_predictions = (full_probabilities >= threshold).astype(int)

```

```

# 6. Save predictions
full_sample["relevance_score"] = full_probabilities
full_sample["predicted_label"] = np.where(full_predictions == 1, "Relevant", "Irrelevant")

# 7. Optional confidence band
def confidence_band(prob):
    if prob >= 0.90 or prob <= 0.10:
        return "high"
    elif prob >= 0.70 or prob <= 0.30:
        return "medium"
    else:
        return "low"

full_sample["confidence"] = full_sample["relevance_score"].apply(confidence_band)

# 8. Save file
full_sample.to_csv("classified_full_sample_first_gpr_type.csv", index=False)

print("Saved: classified_full_sample_first_gpr_type.csv")
print(full_sample["predicted_label"].value_counts(dropna=False))
print(full_sample["confidence"].value_counts(dropna=False))

```

Annex IX: Relevancy Classification Stata Validation Sample Prompt Example – GPR
Group 1

*

=====

* 1. IMPORT DATA

*

=====

```
import excel "C:\Users\Pedro Carvalho\Desktop\Tese\3. Metodologia\II - The Treatment\II - False Positives Removal\Grupo 1 - Novos conflitos globais ou descontrolo dos atuais\classified_full_sample_first_gpr_type.xlsx", firstrow clear
```

* Check variables

```
describe
```

*

=====

* 2. CREATE NUMERIC LABEL (optional but useful)

*

=====

```
gen label_num = .  
replace label_num = 1 if predicted_label == "Relevant"  
replace label_num = 0 if predicted_label == "Irrelevant"
```

*

=====

* 3. RANDOM SEED (for reproducibility)

```

*
=====

set seed 12345

*
=====

* 4. SPLIT DATA BY CLASS

*
=====

* ----- Relevant -----
preserve
keep if predicted_label == "Relevant"

* Randomize
gen rand = runiform()

* Sort by confidence (descending) and random within that
gsort -confidence rand

* Keep 100
keep in 1/100

tempfile relevant_sample
save `relevant_sample'
restore

* ----- Irrelevant -----
preserve
keep if predicted_label == "Irrelevant"

```

```
gen rand = runiform()
```

```
gsort -confidence rand
```

```
keep in 1/100
```

```
tempfile irrelevant_sample
```

```
save `irrelevant_sample'
```

```
restore
```

```
*
```

```
=====
```

```
* 5. COMBINE BOTH SAMPLES
```

```
*
```

```
=====
```

```
use `relevant_sample', clear
```

```
append using `irrelevant_sample'
```

```
* Shuffle final dataset
```

```
gen rand2 = runiform()
```

```
sort rand2
```

```
drop rand rand2
```

```
*
```

```
=====
```

```
* 6. EXPORT FINAL SAMPLE
```

```
*
```

```
=====
```

```
export excel using "validation_sample_200.xlsx", firstrow(variables) replace
```

```
display "Validation sample saved: validation_sample_200.xlsx"
```

Annex X: Relevancy Classification Validation Sample Statistical Details

Group of GPR	Python Classification	Testing Sample
1 - Novos conflitos globais ou descontrolo dos atuais	Relevant	100
	Irrelevant	100
2 - Novos conflitos na Europa ou descontrolo dos atuais	Relevant	100
	Irrelevant	100
3 - Explosão nuclear ou libertação de agentes biológicos ou químicos	Relevant	20
	Irrelevant	20
4 - Ciberataques de grande dimensão a infraestruturas críticas ou empresas	Relevant	30
	Irrelevant	30
5 - Migrações em massa usadas como arma política	Relevant	50
	Irrelevant	50
6 - Radicalização social e política, greves e motins civis violentos	Relevant	100
	Irrelevant	100
7 - Choques energéticos ligados ao petróleo e gás natural	Relevant	20
	Irrelevant	20
8 - Crise financeira grave	Relevant	20
	Irrelevant	20
9 - Disrupção das cadeias de abastecimento globais	Relevant	15
	Irrelevant	85
10 - Conflitos comerciais de elevada intensidade (EUA, China, UE)	Relevant	20
	Irrelevant	130
11 - Desinformação ge-	Relevant	20

rada por Inteligência Artificial	Irrelevant	20
12 - Catástrofes climáticas extremas	Relevant	50
	Irrelevant	100

Annex XI: Framing Classification AI Prompt

Role	/	System	Prompt
You are a research assistant specialized in media framing analysis and geopolitical risk perception, supporting an academic thesis in economics and political economy.			

User	Prompt
You will receive an Excel file where each row corresponds to one news article headline.	
The dataset contains at least the following information per row:	
• Geopolitical Risk Category (already validated as relevant) • Headline (the news headline)	
Your task is to analyze how geopolitical risk is framed in each headline.	

Objective

Classify the framing of geopolitical risk conveyed by each headline, focusing on how the risk is presented, emphasized, and contextualized, rather than on whether the risk objectively exists.

This classification captures subjective risk amplification or attenuation, consistent with the literature on framing effects, availability heuristics, and emotional versus cognitive risk perception.

Conceptual	definition	of	framing
Framing refers to how information about a risk is presented, including:			
• language intensity, • emphasis on threat versus control, • emotional tone, • urgency and severity cues, • narrative focus on escalation or mitigation.			
Framing influences how audiences perceive:			
• likelihood,			

- severity,
- urgency,
- and potential consequences of geopolitical risks.

Framing

categories

For each headline, assign one of the following three categories:

1. Concerning framing

Use this label if the headline:

- Emphasizes threat, danger, escalation, or instability
- Uses alarmist, emotionally charged, or urgent language
- Highlights worst-case scenarios, uncertainty, or lack of control
- Stresses severe economic, political, or security consequences
- Suggests high likelihood or imminent materialization of risk

➡ This framing amplifies perceived geopolitical risk.

2. Neutral framing

Use this label if the headline:

- Presents information in a descriptive, factual, or balanced manner
- Reports events without strong emotional or evaluative language
- Avoids explicit judgments about severity or urgency
- Includes both risks and mitigating factors without emphasis

➡ This framing neither amplifies nor attenuates perceived risk.

3. Dismissive framing

Use this label if the headline:

- Downplays the severity or likelihood of the risk

- Emphasizes control, resilience, adaptation, or normalization
- Frames the risk as manageable, exaggerated, or unlikely
- Focuses on continuity, stability, or successful mitigation

➡ This framing attenuates perceived geopolitical risk.

Important classification rules

- Focus on the overall framing conveyed by the headline.
- Since only the headline is available, base the classification exclusively on the wording and tone of the headline.
- Do not infer details that are not signaled in the headline.
- Do not assess factual accuracy or correctness.
- Do not evaluate whether the risk actually materialized.
- If a headline is ambiguous or could fit two categories, choose **Neutral**.
- Use the dominant signal in the headline: threat amplification, balance, or attenuation.

Instructions

1. Evaluate each headline independently.
2. Do not modify existing columns.
3. Do not remove any observations.
4. Do not include explanations inside the dataset.

Output

format

Return the dataset with one additional column, named exactly:

`framing_classification`

Each row must contain only one of the following values:

- Concerning
- Neutral

- Dismissive

Post-processing**summary**

After the table, provide a short summary including:

- total number of headlines classified,
- number classified as Concerning,
- number classified as Neutral,
- number classified as Dismissive.

I can also tighten this into a more compact version optimized for ChatGPT, Claude, or Gemini prompt input.

Annex XII: Detailed Analysis of the Final Calculations of the MW-GPR Index

The Volume

It can be measured as an objective and/or as a subjective metric, where an objective approach comprehends the exact number of news' articles published in each group of GPR and a subjective approach focuses on the percentage of articles published in one group of GPR in relation to the number of GPR articles published in the other groups. This methodology followed the latter, as the former is a much more abstract concept which would be much harder to compute and translate into a meaningful result.

The Sentiment

Date Coefficient: According to Wahlberg and Sjöberg (2000), the influence of media coverage decreases over time and thus it seems essential to reduce the impact of the articles that were published earlier in the selected period. Therefore, the more recent the articles, the greater sentiment they carry. Additionally, date intervals were used instead of a continuous scale because the impact of media coverage was assumed to decrease gradually over broader periods of time, rather than changing meaningfully from one specific day to the next. This approach also allowed the date coefficient to remain simple, transparent, and comparable across articles.

Outlet Coefficient: According to Jones (2025), the source's trustworthiness positively influences the perception of GPR. Therefore, different outlets can instill different levels of sentiment. The selection of which online newspapers would be included in this study has already taken this into consideration. However, there are still some differences between each outlet, as evidenced by their range and by their target audiences. These differences must be reflected in the coefficients of each outlet.

This coefficient mustn't, however, dominate the model and therefore, the coefficients attributed to each newspaper were kept within a narrow range around 1. Since the final sentiment score is obtained through the multiplication of several coefficients, this approach ensures that the outlet variable influences the result moderately, without disproportionately affecting the overall score.

Expresso and **Público** are two of the biggest entities in Portuguese media. Their combination of extensive range and credibility assures their perception of GPR to be widely spread and transmitted. Therefore, their sentiment must be highlighted, with a slight advantage to

Expresso due to its more global oriented analysis.

Negócios is the most trusted Portuguese economics newspaper. However, its focus on economic issues and limited range positions it slightly behind Expresso and Público.

Observador is the only digital native outlet of the analysis, which, while assuring diversity in the analysis, limits the range and trustworthiness of the newspaper, especially among older audiences. Therefore, it sits slightly behind Negócios.

Finally, **Diário de Notícias** adds the institutional side to the sample. However, its range and relevance have been decaying over the last few years and thus it doesn't stand on the same level as the other 4 outlets.

Framing Coefficient: However important Date and Outlet are, the sentiment expressed in the headlines is much more relevant, as it is the main component that influences the perception of GPR. Therefore, the intervals for this coefficient had to be larger than for the previous variables. As we saw above, the headline could have been classified as "Concerning" (if the headline actively increases the perception of GPR), "Neutral" (if the headline is neutral or doesn't influence the perception of GPR) or "Dismissive" (if the headline actively decreases the perception of GPR).

The Composite

The results of the sensitivity analysis can be found in Annex XVII.

The Abnormal Sentiment

According to Venade and Grilo (2023), Abnormal Sentiment captures the extent to which the sentiment associated with each GPR group differs from its expected level. While the sentiment indicator measures the overall tone of the articles, abnormal sentiment focuses on the unexpected component of that tone, identifying whether a group presents a sentiment value above or below what would normally be anticipated. To calculate it, the authors measured the difference between the observed sentiment and its benchmark value. This difference was then converted into a 1 to 5 scale to allow comparison across GPR groups. Higher values therefore indicate stronger positive abnormal sentiment, while lower values reflect weaker or negative abnormal sentiment (Venade and Grilo, 2023).

Although this dissertation followed in the footsteps of Venade and Grilo (2023), some adjustments were introduced to better adapt the analysis to the characteristics of the data base,

the objectives of the study and the specific structure of the indicators being compared. The specific methodology followed was as explained below.

Firstly, although Venade and Grilo's model considered solely the sentiment, it seems relevant in this study to consider both the Sentiment and the Volume in a Composite way, as media relevance of each GPR group depends not only on the tone expressed in the articles, but also on how frequently that topic appears in the news sample. Therefore, despite being called Abnormal Sentiment it will consider the Composite variable that was analyzed above.

Then, in Venade and Grilo's model the abnormal sentiment was measured by comparing the sentiment with a benchmark. In this study, however, there is no benchmark. Therefore, the Abnormal Sentiment will be measured by comparing the difference between the actual Composite value of the most recent quarter and the expected Composite value of the most recent quarter, computed as a linear regression based on the first three quarters. This difference was then transformed into a 1 to 5 scale.

Annex XIII: Initial MW-GPR Index Database

Group of GPR	Keyword	Newspapers					Total	
		Expresso	Negócios	Público	Observador	Diário Notícias	by Keyword	by Group of GPR
1 - Novos conflitos globais ou descontrolo dos atuais	1	264	24	213	170	219	890	2048
	2	94	25	197	190	83	589	
	3	189	40	63	164	113	569	
2 - Novos conflitos na Europa ou descontrolo dos atuais	1	234	45	244	246	215	984	1670
	2	99	4	48	31	25	207	
	3	165	40	52	137	85	479	
3 - Explosão nuclear ou libertação de agentes biológicos ou químicos	1	97	1	7	17	9	131	409
	2	40	7	35	27	15	124	
	3	81	8	21	27	17	154	
4 - Ciberataques de grande dimensão a infraestruturas críticas ou empresas	1	81	6	17	25	30	159	583
	2	36	94	93	24	6	253	
	3	51	20	43	31	26	171	
5 - Migrações em massa usadas como arma política	1	238	276	235	256	266	1271	1825
	2	77	33	105	62	117	394	
	3	50	5	47	32	26	160	
6 - Radicalização social e política, greves e motins civis violentos	1	232	28	233	236	257	986	2777
	2	239	84	225	222	219	989	
	3	253	20	193	172	164	802	
7 - Choques energéticos ligados ao petróleo e gás natural	1	114	139	47	69	33	402	755
	2	51	88	81	41	38	299	
	3	25	4	6	15	4	54	
8 - Crise financeira grave	1	69	64	129	84	90	436	564
	2	14	18	13	22	16	83	
	3	13	7	11	4	10	45	
9 - Disrupção das cadeias de abastecimento globais	1	14	33	19	15	27	108	872
	2	8	20	58	36	23	145	
	3	23	217	247	117	15	619	
10 - Conflitos comerciais de elevada intensidade (EUA, China, UE)	1	18	39	39	7	10	113	1758
	2	251	231	222	117	101	922	
	3	320	268	39	68	28	723	
11 - Desinformação gerada por Inteligência Artificial	1	1	1	2	5	2	11	419
	2	37	3	74	88	57	259	
	3	16	18	39	52	24	149	
12 - Catástrofes climáticas extremas	1	14	2	79	49	14	158	1600
	2	68	14	165	84	70	401	
	3	190	295	217	176	163	1041	
Total		3766	2221	3558	3118	2617	15280	

Annex XIV: Opinion Piece Removal Testing Sample

		Manual Classification	
		Not Opinion Piece	Opinion Piece
AI Classification	Not Opinion Piece	288	12
	Opinion Piece	4	296

Not Opinion Piece – Not Opinion Piece	96%
Opinion Piece – Opinion Piece	99%

Annex XV: MW-GPR Index Database After Removal of Opinion Pieces

Group of GPR	Keyword	Newspapers					Total	
		Expresso	Negócios	Público	Observador	Diário Notícias	by Keyword	by Group of GPR
1 - Novos conflitos globais ou descontrolo dos atuais	1	260	22	206	146	183	817	1893
	2	90	25	194	172	73	554	
	3	186	36	61	144	95	522	
2 - Novos conflitos na Europa ou descontrolo dos atuais	1	233	41	235	213	176	898	1537
	2	98	4	47	25	23	197	
	3	164	37	49	120	72	442	
3 - Explosão nuclear ou libertação de agentes biológicos ou químicos	1	95	1	7	12	6	121	376
	2	38	7	32	23	13	113	
	3	78	8	20	25	11	142	
4 - Ciberataques de grande dimensão a infraestruturas críticas ou empresas	1	80	5	16	25	27	153	564
	2	35	93	88	20	6	242	
	3	50	19	43	31	26	169	
5 - Migrações em massa usadas como arma política	1	225	265	228	202	244	1164	1674
	2	74	32	103	50	105	364	
	3	48	5	45	25	23	146	
6 - Radicalização social e política, greves e motins civis violentos	1	218	27	226	193	244	908	2486
	2	233	80	214	180	215	922	
	3	232	20	177	102	125	656	
7 - Choques energéticos ligados ao petróleo e gás natural	1	110	139	47	61	29	386	703
	2	50	84	78	22	33	267	
	3	25	4	6	11	4	50	
8 - Crise financeira grave	1	61	62	124	60	80	387	505
	2	13	17	13	16	15	74	
	3	13	7	11	3	10	44	
9 - Disrupção das cadeias de abastecimento globais	1	14	33	19	12	19	97	838
	2	7	20	58	29	21	135	
	3	22	213	243	117	11	606	
10 - Conflitos comerciais de elevada intensidade (EUA, China, UE)	1	18	38	36	3	6	101	1665
	2	251	222	212	99	92	876	
	3	318	258	39	58	15	688	
11 - Desinformação gerada por Inteligência Artificial	1	1	1	2	4	0	8	328
	2	34	3	66	50	45	198	
	3	13	18	36	36	19	122	
12 - Catástrofes climáticas extremas	1	14	2	79	39	11	145	1470
	2	67	14	159	74	66	380	
	3	176	293	210	131	135	945	
Total		3644	2155	3429	2533	2278	14039	

Annex XVI: Irrelevant Articles Removal Testing Samples

1 - Novos conflitos globais ou descontrolo dos atuais			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	94	5
		Irrelevant	6	95
	Relevant - Relevant			94%
2 - Novos conflitos na Europa ou descontrolo dos atuais			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	90	1
		Irrelevant	10	99
	Relevant - Relevant			90%
	Irrelevant - Irrelevant			99%
3 - Explosão nuclear ou libertação de agentes biológicos ou químicos			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	18	0
		Irrelevant	2	20
	Relevant - Relevant			90%
4 - Ciberataques de grande dimensão a infraestruturas críticas ou empresas			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	24	0
		Irrelevant	6	30
	Relevant - Relevant			80%
5 - Migrações em massa			Manual Classification	
			Relevant	Irrelevant

usadas como arma política	Python Classification	Relevant	43	3
		Irrelevant	7	47
	Relevant – Relevant			86%
	Irrelevant – Irrelevant			94%
6 - Radicalização social e política, greves e motins civis violentos			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	90	4
		Irrelevant	10	96
	Relevant - Relevant			90%
	Irrelevant - Irrelevant			96%
7 - Choques energéticos ligados ao petróleo e gás natural			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	18	0
		Irrelevant	2	20
	Relevant - Relevant			90%
	Irrelevant - Irrelevant			100%
8 - Crise financeira grave			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	18	0
		Irrelevant	2	20
	Relevant - Relevant			90%
	Irrelevant - Irrelevant			100%
9 - Disrupção das cadeias de abastecimento globais			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	13	1
		Irrelevant	2	84
	Relevant - Relevant			87%

	Irrelevant - Irrelevant			99%
10 - Conflitos comerciais de elevada intensidade (EUA, China, UE)			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	18	1
		Irrelevant	2	129
	Relevant - Relevant			90%
	Irrelevant - Irrelevant			99%
11 - Desinformação gerada por Inteligência Artificial			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	17	1
		Irrelevant	3	19
	Relevant - Relevant			85%
	Irrelevant - Irrelevant			95%
12 - Catástrofes climáticas extremas			Manual Classification	
			Relevant	Irrelevant
	Python Classification	Relevant	46	3
		Irrelevant	4	97
	Relevant - Relevant			92%
	Irrelevant - Irrelevant			97%

Annex XVII: Sensitivity Analysis I

Sentiment			Short Term								Long Term							
			Short Term	Micro-Emp	PME	GrandEmp	OS = 0	OS = 1	OS = 2	OS = 3	Long Term	Micro-Emp	PME	GrandEmp	OS = 0	OS = 1	OS = 2	OS = 3
C- DATE	1	1,1	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		1,05	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		1	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,95	43,36%	29,77%	20,67%	56,39%	30,82%	30,69%	24,52%	51,05%	27,27%	5,59%	16,84%	51,05%	22,38%	6,33%	3,85%	60,60%
		0,9	39,16%	24,17%	21,02%	49,39%	28,37%	23,28%	20,32%	52,45%	20,28%	-1,40%	14,04%	45,45%	17,48%	-2,81%	1,40%	58,49%
	0,75	0,825	44,06%	28,72%	25,22%	54,99%	31,87%	31,75%	26,97%	49,65%	28,67%	6,29%	20,35%	48,95%	23,08%	7,03%	9,81%	63,75%
		0,788	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,75	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,713	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,675	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%

	0,5	0,55	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,525	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,5	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,475	46,85%	32,22%	24,17%	60,95%	34,68%	35,63%	27,67%	52,45%	32,17%	9,79%	20,70%	52,45%	25,17%	11,25%	8,41%	63,05%
		0,45	46,85%	32,22%	24,17%	60,95%	34,68%	35,63%	27,67%	52,45%	32,17%	9,79%	20,70%	52,45%	25,17%	11,25%	8,41%	63,05%
	0,25	0,275	46,85%	32,22%	24,17%	60,95%	34,68%	35,63%	27,67%	52,45%	32,17%	9,79%	20,70%	52,45%	25,17%	11,25%	8,41%	63,05%
		0,263	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,25	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,238	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,225	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%

C-PAPER	1,2	1,32	40,56%	24,87%	21,72%	51,84%	30,12%	25,40%	21,72%	54,55%	21,68%	-0,70%	15,09%	46,15%	18,18%	-1,76%	3,15%	59,19%
		1,26	44,76%	30,47%	21,37%	58,84%	32,57%	32,81%	25,92%	53,15%	28,67%	6,29%	17,89%	51,75%	23,08%	7,38%	5,60%	61,30%
		1,2	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		1,14	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		1,08	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	1,1	1,21	45,45%	29,42%	25,92%	57,44%	33,63%	33,86%	28,37%	51,75%	30,07%	6,99%	21,40%	49,65%	23,78%	8,08%	11,56%	64,45%
		1,155	44,06%	28,72%	25,22%	54,99%	31,87%	31,75%	26,97%	49,65%	28,67%	6,29%	20,35%	48,95%	23,08%	7,03%	9,81%	63,75%
		1,1	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		1,045	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,99	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%

1	1,1	42,66%	28,02%	18,91%	59,19%	28,02%	32,10%	26,97%	46,85%	29,37%	6,29%	16,49%	53,15%	20,98%	9,14%	4,20%	59,90%
	1,05	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	1	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	0,95	45,45%	29,42%	25,92%	57,44%	33,63%	33,86%	28,37%	51,75%	30,07%	6,99%	21,40%	49,65%	23,78%	8,08%	11,56%	64,45%
	0,9	37,76%	23,47%	20,67%	47,29%	29,07%	21,87%	19,61%	52,45%	17,48%	-4,20%	12,63%	41,26%	15,38%	-5,98%	3,15%	56,39%
	0,99	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	0,945	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	0,9	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	0,855	46,85%	32,22%	24,17%	60,95%	34,68%	35,63%	27,67%	52,45%	32,17%	9,79%	20,70%	52,45%	25,17%	11,25%	8,41%	63,05%
	0,81	46,85%	32,22%	24,17%	60,95%	34,68%	35,63%	27,67%	52,45%	32,17%	9,79%	20,70%	52,45%	25,17%	11,25%	8,41%	63,05%

	0,8	0,88	46,85%	32,22%	24,17%	60,95%	34,68%	35,63%	27,67%	52,45%	32,17%	9,79%	20,70%	52,45%	25,17%	11,25%	8,41%	63,05%
		0,84	46,85%	32,22%	24,17%	60,95%	34,68%	35,63%	27,67%	52,45%	32,17%	9,79%	20,70%	52,45%	25,17%	11,25%	8,41%	63,05%
		0,8	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,76	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		0,72	43,36%	29,77%	20,67%	56,39%	30,82%	30,69%	24,52%	51,05%	27,27%	5,59%	16,84%	51,05%	22,38%	6,33%	3,85%	60,60%
C-FRAMING	1,5	1,65	41,96%	27,67%	19,97%	55,34%	31,17%	27,16%	21,02%	55,24%	23,78%	2,10%	14,39%	48,95%	19,58%	1,41%	0,00%	57,79%
		1,575	43,36%	29,77%	20,67%	56,39%	30,82%	30,69%	24,52%	51,05%	27,27%	5,59%	16,84%	51,05%	22,38%	6,33%	3,85%	60,60%
		1,5	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		1,425	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
		1,35	44,06%	28,72%	25,22%	54,99%	31,87%	31,75%	26,97%	49,65%	28,67%	6,29%	20,35%	48,95%	23,08%	7,03%	9,81%	63,75%

1	1,1	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	1,05	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	1	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	0,95	43,36%	29,77%	20,67%	56,39%	30,82%	30,69%	24,52%	51,05%	27,27%	5,59%	16,84%	51,05%	22,38%	6,33%	3,85%	60,60%
	0,9	41,96%	27,67%	19,97%	55,34%	31,17%	27,16%	21,02%	55,24%	23,78%	2,10%	14,39%	48,95%	19,58%	1,41%	0,00%	57,79%
0,5	0,55	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	0,525	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	0,5	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	0,475	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%
	0,45	45,45%	31,52%	23,47%	58,49%	32,92%	33,51%	26,27%	50,35%	30,77%	9,09%	19,65%	51,75%	24,48%	10,19%	6,66%	62,35%

Annex XVIII: Sensitivity Analysis II

Composite			Short Term							
			Short Term	MicroEmp	PME	GrandEmp	OS = 0	OS = 1	OS = 2	OS = 3
Sentiment	0,5	0,700	0,00%	-10,86%	-9,11%	15,41%	-3,85%	-17,28%	-11,21%	27,27%
		0,600	0,70%	-9,11%	-10,51%	16,46%	-3,85%	-18,34%	-9,46%	30,07%
		0,500	4,20%	-8,41%	-1,40%	17,16%	-1,05%	-12,70%	-3,85%	37,06%
		0,400	0,70%	-11,21%	-4,90%	15,41%	-4,90%	-12,70%	-4,55%	32,17%
		0,300	-6,99%	-17,51%	-10,51%	7,71%	-11,91%	-20,46%	-11,56%	27,27%

Composite			Long Term							
			Long Term	MicroEmp	PME	GrandEmp	OS = 0	OS = 1	OS = 2	OS = 3
Sentiment	0,5	0,700	-22,38%	-42,66%	-20,70%	1,40%	-23,08%	-45,34%	-25,57%	16,46%
		0,600	-24,48%	-45,45%	-22,46%	1,40%	-23,08%	-47,10%	-27,32%	14,01%
		0,500	-19,58%	-40,56%	-11,58%	0,70%	-16,78%	-42,18%	-18,21%	17,86%
		0,400	-20,98%	-41,96%	-12,63%	-3,50%	-18,88%	-42,18%	-17,16%	14,36%
		0,300	-28,67%	-48,25%	-18,25%	-11,19%	-25,17%	-49,21%	-23,47%	6,66%

FACULDADE DE ECONOMIA

