Speech and Standing:

The Impact of Economic Fluctuations and Global Shocks on Rhetoric in the UN General Debate

Abstract

The United Nations General Debate provides a longitudinal record of how states articulate priorities, diagnose problems, and signal cooperation before a global audience. Because speeches are not negotiated texts but individual statements of preference, they offer insight into both national positioning and the collective international order. This study combines supervised and unsupervised methods to examine how discourse evolves over time, varies across countries, and responds to economic conditions and systemic shocks. The analysis evaluates the detectability of rhetoric tied to domestic performance, traces thematic shifts across crises, and assesses whether rhetorical emphasis differs systematically across income groups. Results show that downturn language is diffuse and harder to identify than upturn language, while global shocks compress the agenda around crisis-specific concerns without displacing baseline issues such as security, development, and international law. Cross-national comparisons highlight persistent divides, with higher income states more often stressing security and institutional order, and lower income states emphasizing development and climate vulnerability. These results confirm that the General Debate functions as a barometer of international priorities, reflecting both the economic cycles and the inequalities of the global system.

1. Introduction

The United Nations General Debate (UNGD) is a rare forum where every member state annually articulates its view of international affairs, producing a continuous archive of priorities, grievances, and self-presentation. Because speeches are delivered in a consistent format the corpus provides an exceptional foundation for examining how international discourse evolves over time, differs across countries, and responds to global disruptions.

Patterns of rhetoric within this setting shift along two dimensions. The first reflects variation in issue framing, as themes emerge, recede, or diverge across states and across time. Examining these shifts highlights both the evolution of the UN agenda and the degree to which countries converge on shared priorities or advance distinct narratives. The second dimension reflects the influence of economic context. Domestic conditions such as growth and contraction shape rhetorical tone, while systemic shocks such as wars, pandemics, or financial crises impose collective stress that narrows attention and reorders agendas.

To capture these dynamics, the study applies a mixed-methods text mining strategy. Supervised classification assesses whether economic regimes can be reliably detected in speech texts, while unsupervised models trace the thematic structures that characterize debate during stable periods and under crisis conditions. Cross-national comparisons link rhetorical variation to income

groups, testing whether structural inequality shapes the emphasis countries place on themes such as security, development, and climate. Together these approaches provide a comprehensive view of how the General Debate reflects both gradual transformations in the international order and acute moments of global disruption.

2. Related Work

Research on sentiment and text classification has established the value of supervised models for detecting evaluative tone in large corpora (Pang, Lee, and Vaithyanathan 79). Foundational work demonstrates that simple representations combined with linear classifiers can recover stable patterns in language usage across domains (Joachims 133–44). Topic modeling has also been widely applied to uncover latent themes in political discourse, policy documents, and legislative speech, enabling comparison across time and contexts (Blei, Ng, and Jordan 996–1002; Grimmer and Stewart 267–82). Studies of the UN General Debate have shown that the corpus provides insight into state preferences and international alignment, since leaders are able to present issues in their preferred framing (Baturo, Dasandi, and Mikhaylov 372–86).

Supervised sentiment analysis methods are vital for assessing large-scale text, particularly political language that carries varied tone and emphasis. Pang and Lee show that even simple feature representations, when combined with machine learning classifiers, can accurately detect opinions across domains (Pang and Lee 164–65). Their work demonstrates how computational methods can reliably identify evaluative content, making them especially useful for analyzing how political actors frame arguments and positions in speeches and debates.

Applications of the UN General Debate Corpus further reinforce its value for analyzing international discourse. Baturo, Dasandi, and Mikhaylov introduced the dataset and demonstrated its ability to reveal consistent cross-national patterns in agenda-setting and rhetorical alignment (372–86). Subsequent work has used the corpus to examine state preferences, norm diffusion, and how global events shape issue salience (Mikhaylov, Baturo, and Dasandi 152–70; Mikaylov et al. 405–24). Together these studies show that the UNGD is a reliable source for studying how governments articulate both national and collective priorities.

3. Data

The UN General Debate Corpus was compiled and released through the Harvard Dataverse. It contains annual speeches delivered by member states ("United Nations General Debate Corpus 1946-2024"). The corpus was limited to the years 2005 through 2024 and speech texts were aligned with country codes, year, and speaker identity and were restricted to countries with a complete twenty-year record to maintain comparability across time. National context was added from the World Bank through GDP per capita and income group classifications.

Exploratory assessment shows that speeches cluster around 1,500 to 2,500 words with a right tail of extended addresses. GDP per capita is highly skewed with large dispersion across income groups (Appendix A.1). Speech length is broadly consistent across income groups over time, with a

clear decline during 2020 and subsequent recovery (Appendix A.2). Speaker turnover varies considerably across countries and regions, reflecting differences in political transitions and representation practices (Appendix A.3). Sentiment scores from a lexicon-based tool are overwhelmingly positive, which is consistent with formal diplomatic style and limits the instrument's utility for absolute comparisons; relative comparisons across groups and periods remain informative (Appendix A.4).

4. Methods

4.1 Economic regimes
A supervised classification framework was used to test whether speeches given during economic downturns can be separated from those delivered in upturns. To keep the focus on meaningful content, a custom stopword list was applied to remove ceremonial and institutional vocabulary that was common across all speeches but not tied to economic context. This included terms such as president, assembly, and anniversary numbers that reflect the setting rather than economic conditions. The speeches were then vectorized with TF-IDF, using both unigrams and bigrams. This allowed the model to capture not only individual words but also paired phrases like economic growth or financial crisis, which often carry stronger signals. Frequency thresholds filtered out words that were either too rare or too common, and sublinear weighting reduced the impact of repetition

The model used was a linear Support Vector Machine with balanced class weights to handle the uneven distribution between downturn and upturn speeches. Countries were grouped during training and testing to prevent overlap and ensure the evaluation reflected general patterns rather than country specific phrasing. Hyperparameters for both the vectorizer and the penalty term were tuned with grid search, using macro F1 as the main scoring metric to balance performance across the two classes. Results were reported on held out test countries, with accuracy, precision, recall, and confusion matrices used for evaluation. In addition, thresholds on the decision scores were varied to examine the trade off between precision and recall for downturns, giving a more complete view of how well the model could capture minority class signals.

4.2 Global shocks
Periods of major global shocks were isolated by defining windows for the financial crisis (2008 – 2010), geopolitical conflicts (2015 – 2016), and the pandemic (2020 -2021). Restricting speeches created focused subsets that enable direct comparison across crises. This design makes it possible to trace how discourse changes under conditions of global disruption. On each window, topic models were trained separately to uncover the thematic structure of discourse. Speeches were cleaned, tokenized, and processed into candidate models with varying numbers of topics. Coherence values guided selection of the best-fitting model in each case, ensuring that topics were both interpretable and distinct. This approach is well suited to the UN General Debate corpus, as it reduces a large and diverse set of political texts into a manageable set of themes without imposing predefined categories. By modeling each shock period independently, the analysis captures how thematic emphasis shifts under crisis conditions, while still allowing for continuity in long standing concerns.

To support interpretation, topics were labeled at two complementary levels of detail. Granular labels captured the specific contexts and issues, refugee movements, vaccine equity, or regional conflicts emerging in each period. Broad labels consolidated these into higher order categories like "Wars & Security" or "Global Health & Solidarity," enabling comparison across shocks without losing interpretability. The resulting topic summaries provide a standardized format that combines mean prevalence, representative keywords, and thematic labels, offering a structured basis for examining how crises reshape the UN agenda.

5. Results

5.1 Economic regimes: detectability and language cues
The supervised classification model reached an overall accuracy of 0.81 when distinguishing between downturn and upturn speeches. Performance differed across the two classes due to the imbalance (Appendix A.5). Upturn speeches were identified with high precision and recall, while downturn speeches had lower recall at 0.35. The macro F1 score of 0.65 reflected this disparity by weighting both classes equally. Confusion matrices indicated that downturn speeches were more often misclassified as upturns, showing uneven detectability between the two economic regimes.

To address the imbalance, thresholds on the SVM decision scores were adjusted to examine the precision–recall trade-off for downturns. This tuning increased recall for downturn speeches from 0.35 to 0.41 while maintaining precision at 0.45. Overall accuracy declined slightly to 0.78. The revised confusion matrix showed a higher share of downturns correctly identified while upturn classification remained stable, with precision and recall both above 0.85 (Appendix A.6).

Feature weights from a bigram-only vectorization showed distinct linguistic markers for each regime. Downturn speeches frequently included terms such as economic crisis, covid pandemic, and financial crisis. Upturn speeches more often contained phrases such as economic growth, development agenda, and regional cooperation. These paired terms provided contextual cues that differentiated the two categories (Appendix A.7).

5.2 Global shocks: thematic reconfiguration and baseline contrast
Average sentiment in General Debate speeches declined during global shocks, with dips observed during the financial crisis, geopolitical conflicts, and COVID-19 (Appendix A.8). Sentiment generally returned to prior levels after each event.

Topic modeling indicated that each crisis produced distinct themes, with vocabulary and coherence values marking how these themes were expressed (Appendix A.9). The financial crisis emphasized food security, economic instability, and terrorism, with coherence peaking at six topics. Geopolitical conflicts featured refugees, sovereignty disputes, and regional crises such as Syria and Ukraine, showing the clearest separation with coherence peaking at eight topics. COVID-19 reordered the agenda toward vaccines, global solidarity, and public health, reflecting overlapping concerns during a global emergency. Across all crises, themes related to security and international law remained present (Appendix A.10).

Comparison with non-shock periods showed broader thematic distributions in calmer years, including democracy, governance, climate change, and development. During crises, themes were more concentrated. The financial crisis emphasized economic stability and security, geopolitical conflicts centered on refugees and sovereignty, and COVID-19 focused on health and survival. These results highlight differences in the breadth of issues across stable versus crisis conditions (Appendix A.11).

6. Discussion

Economic Conditions and Shocks
The analysis of economic regimes revealed an asymmetry in how states communicate during periods of growth versus contraction. Upturns are framed in consistent, cooperative language that models could detect with high reliability. Downturns, by contrast, were less systematically articulated. Even after threshold adjustments improved recall, downturn speeches remained diffuse and harder to capture. This suggests that while positive economic contexts invite a common vocabulary of growth and stability, negative conditions are expressed through more heterogeneous rhetorical strategies, often couched in aspirational or diplomatic phrasing that masks distress. The challenge of detecting downturns underscores the limits of surface level linguistic signals in capturing economic vulnerability within formal political discourse.

Global shocks offered a different lens by isolating moments of collective disruption. Here, rhetorical shifts were more pronounced and more readily detectable. Sentiment declined across crises, with vocabulary and topic modeling showing that each event left a distinct thematic imprint: economic security in the financial crisis, displacement and sovereignty in the geopolitical conflict period, and health and solidarity during COVID-19. These shifts demonstrate that shocks compress the agenda, narrowing attention to the most immediate threats while temporarily displacing broader concerns. Non-shock years, by contrast, maintained a wider thematic spread, highlighting how the debate normally accommodates multiple competing priorities.

Together, these findings show that economic conditions shape discourse in both subtle and acute ways. Upturns and downturns reflect the difficulty of capturing domestic pressures within diplomatic speech, while global shocks leave sharper and more cohesive traces in the language of the debate. The comparison between regime level variation and systemic crises illustrates how discourse adapts across scales, from national economic cycles to global emergencies. This dual perspective emphasizes that the UN General Debate functions not only as a venue for articulating standing priorities but also as a responsive archive of how international rhetoric reorients under global stress.

7. Limitations and robustness

The corpus captures public framing rather than private negotiation, which may favor aspirational language and increase the difficulty of distinguishing negative from positive tone (Mikhaylov, Baturo, and Dasandi 156). Lexicon-based sentiment methods are biased toward positivity in diplomatic speech and are therefore more useful for relative rather than absolute inference (Young

and Soroka 209). Topic labeling also requires judgment and should be supported by top-word lists and representative excerpts, as emphasized in prior work on content analysis (Grimmer and Stewart 270). Income classifications change slowly, and GDP per capita serves as a proxy for economic standing but does not capture distributional or informal dynamics, limiting the depth of socioeconomic interpretation (World Bank). Results are robust to reasonable changes in topic number and vectorization choices, but future work could extend to multilingual analysis and dynamic topic models (Roberts et al. 1065).

Speeches in the General Debate often include euphemistic language when referring to negative or sensitive topics. Leaders may reframe or downplay issues in ways that complicate automated classification, since models rely on consistent cues (Charteris-Black 31). This practice is especially common in high-stakes diplomatic contexts where rhetoric is used strategically.

Another limitation is translation. The corpus consists of English texts, but many speeches were originally delivered in other languages. This introduces the risk that culturally specific tones or rhetorical nuances are lost or flattened during translation (Van Dijk 25). Analyses may therefore unintentionally smooth over linguistic variation, especially for non-English-speaking states.

8. Conclusion

The analysis of two decades of United Nations General Debate speeches shows how international discourse is structured by both cyclical economic conditions and sudden global disruptions. At the domestic level, economic upturns and downturns are not equally visible in rhetoric. Upturns produce a consistent vocabulary of growth and cooperation, while downturns are expressed in more diffuse and aspirational language that is harder to detect systematically. At the systemic level, global shocks leave sharper imprints. The financial crisis, geopolitical conflicts, and the COVID-19 pandemic each narrowed the agenda around urgent concerns, reorienting debate without fully displacing long-standing themes such as peace, development, and multilateralism.

Beyond temporal dynamics, the results underscore persistent rhetorical divides across national income groups. Wealthier states devote more attention to security, institutional reform, and geopolitical competition, while lower-income states emphasize development, climate vulnerability, and regional stability. These patterns highlight how structural inequalities shape not only material capacities but also the language of multilateral diplomacy. Cross-national variation thus complements the crisis analysis by revealing that discourse reflects enduring hierarchies even as it adapts to moments of disruption.

Taken together, the findings confirm that the UN General Debate operates as more than a ceremonial exchange of statements. It functions as a barometer of international priorities, one that registers both the gradual rhythms of economic cycles and the acute shocks of global crises. By tracing how states frame challenges, signal cooperation, and express vulnerability, the corpus offers a valuable record of how the international system interprets and responds to change. This dual perspective—on both structural divides and collective disruptions—demonstrates the value of computational text analysis for capturing the evolving architecture of global discourse.

Appendix A

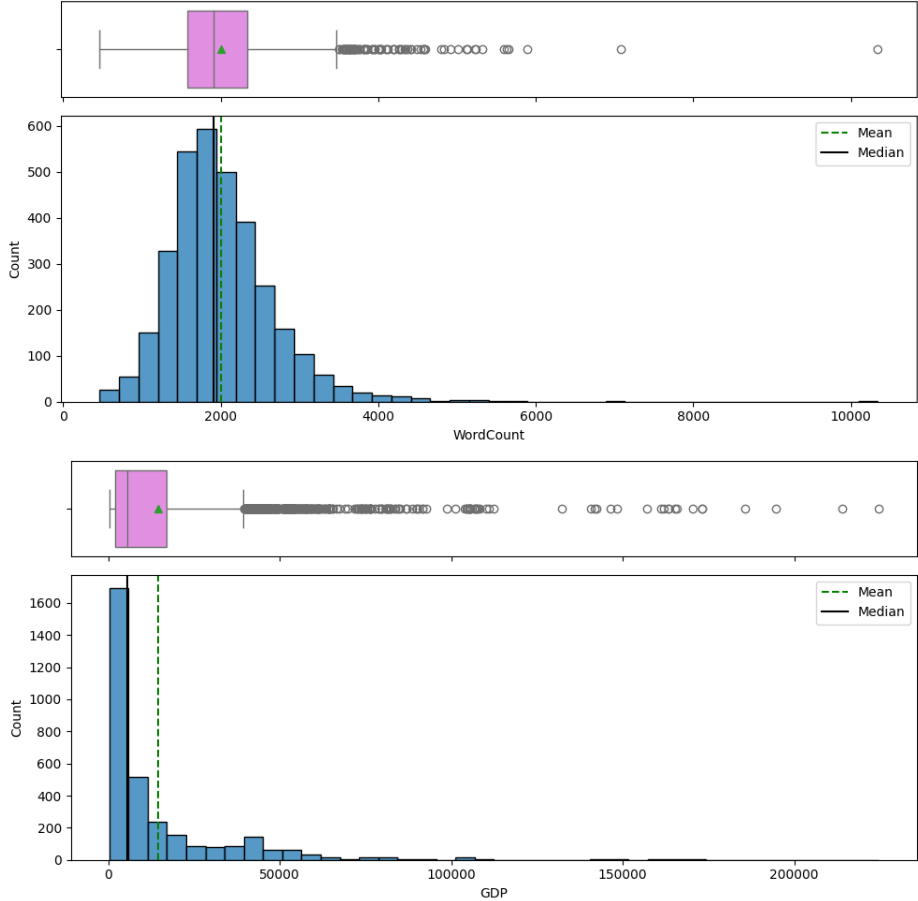**Figure A1:** *Count vs Wordcount & Count vs GDP*

Figure A2: *Average Speech Length Over Time by Income Group*

Average Speech Length Over Time by Income Group

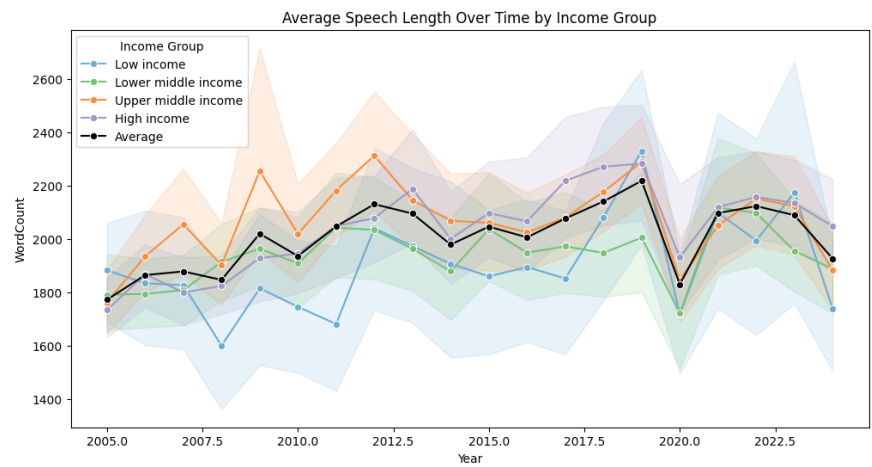**Figure A3:** *Countries Represented by Number of Unique Speakers*

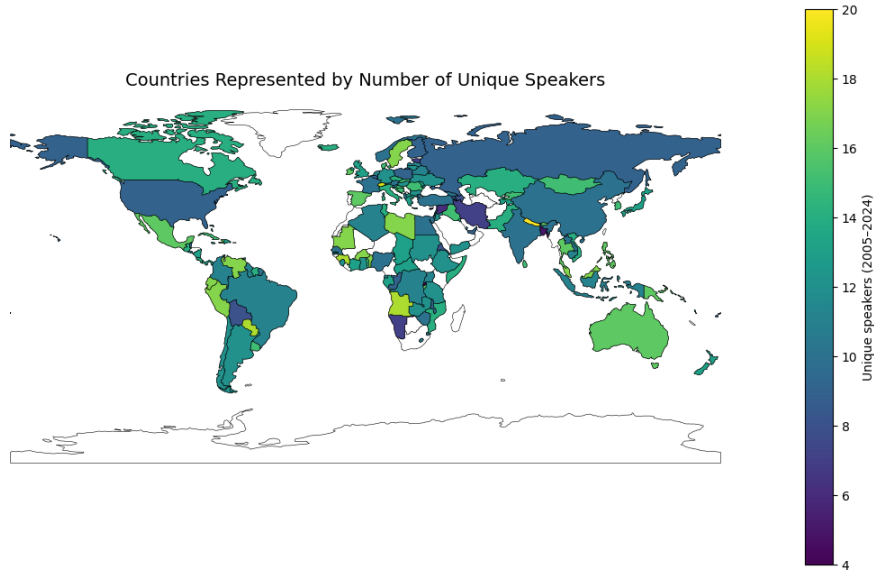Countries Represented by Number of Unique Speakers

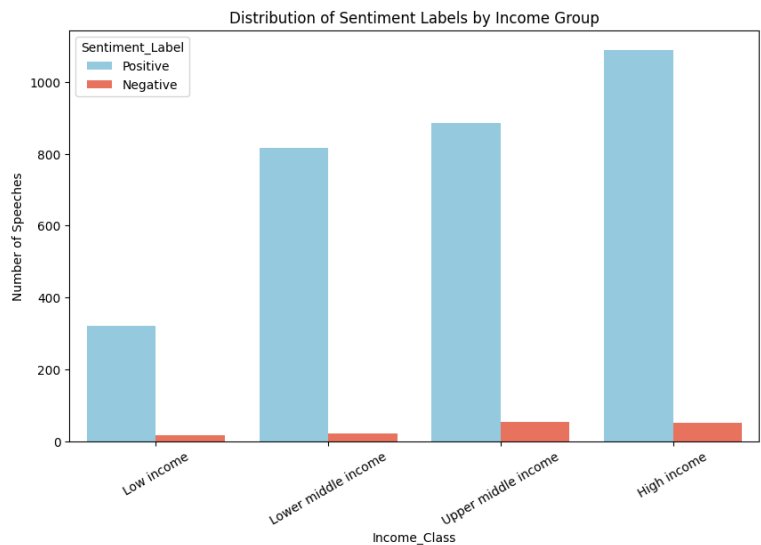**Figure A4:** *Distribution of Sentiment Labels by Income Group*

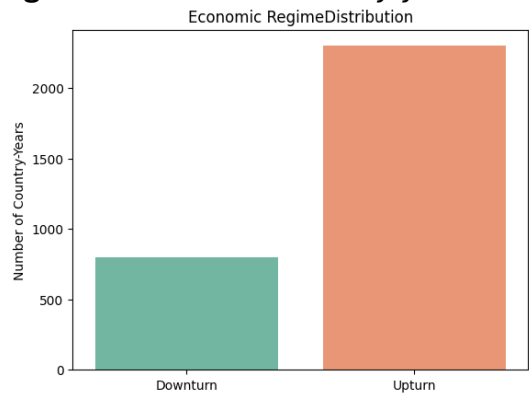**Figure A5:** *Numer of Country-years vs Economic RegimeDistribution*
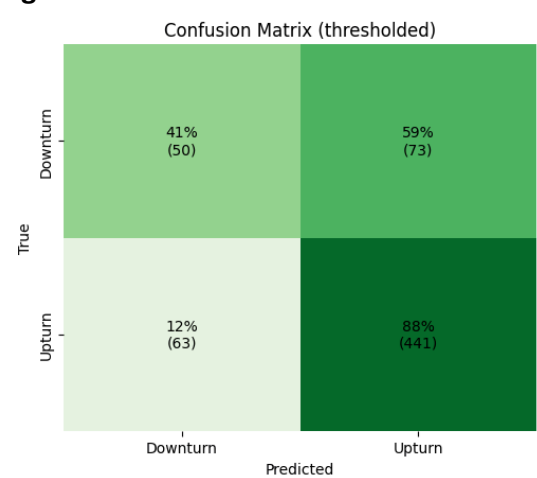

Figure A6: *Confusion Matrix: True vs Predicted*

Figure A7: *Downturn Cues: Term vs Weight, Upturn Cues: Term vs Weight*
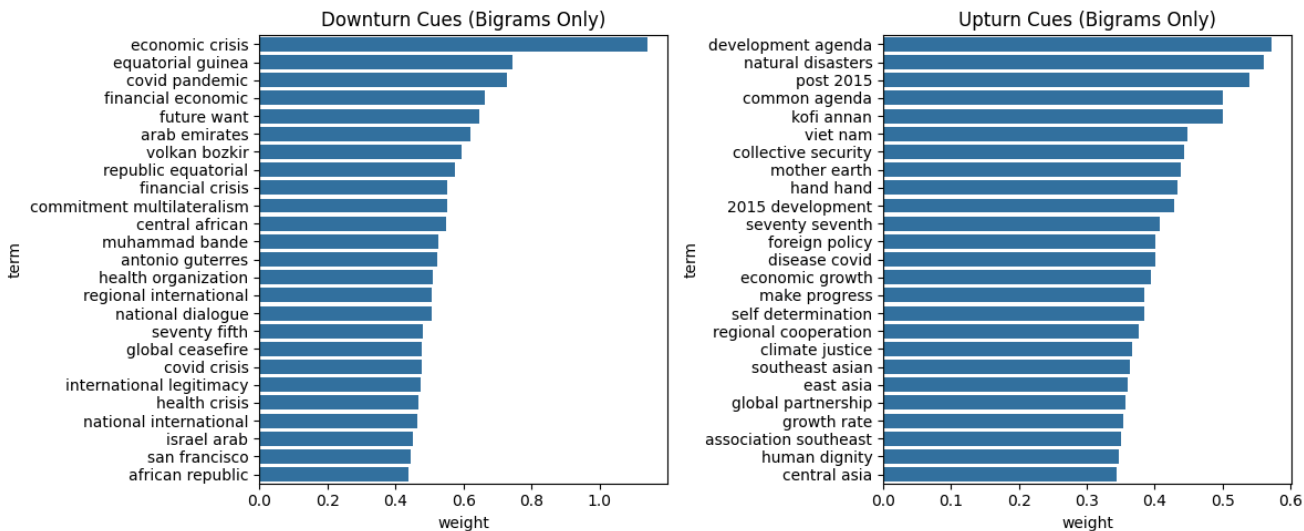
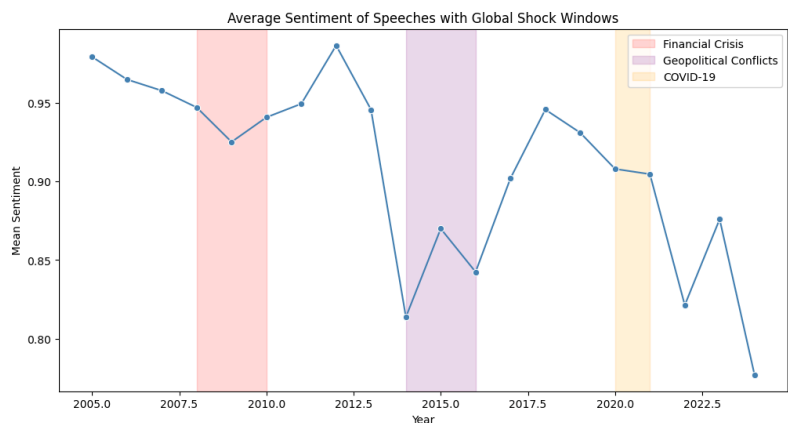**Figure A8:** *Average Sentiment of Speeches with Global Shock Windows*

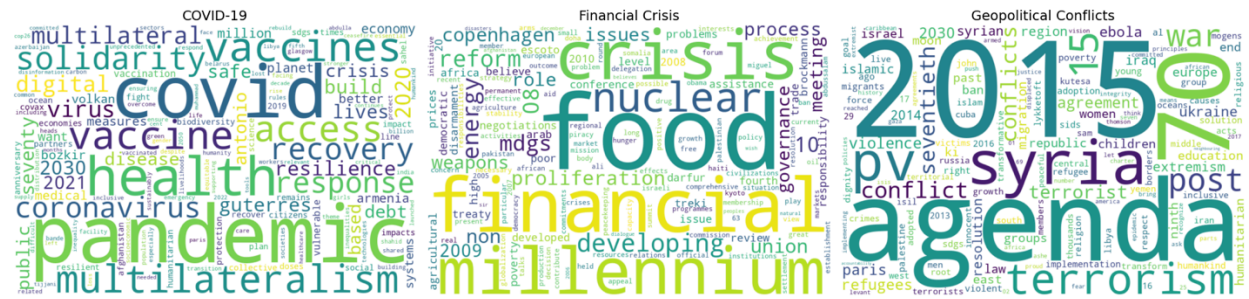**Figure A9:** *Word Cloud: Covid-19, Financial Crisis, & Geopolitical Conflicts*

Figure A10: *Theme Share by Window*

Theme share by window

Figure A11: Theme Prevalence by Window (Broad Categories)
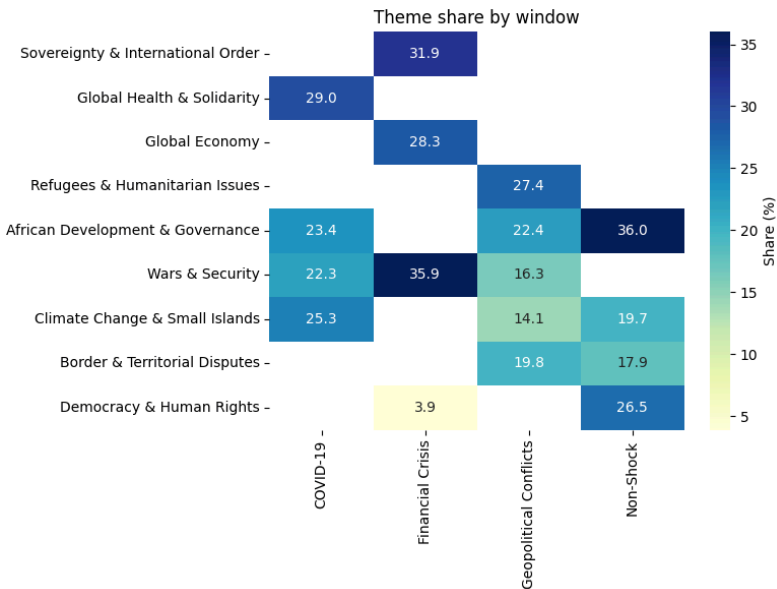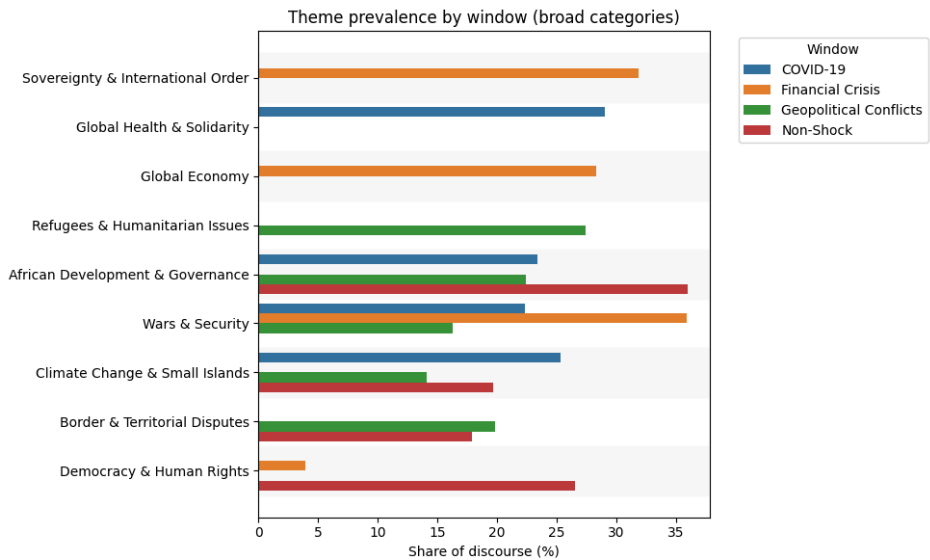
Theme prevalence by window (broad categories)

Works Cited

Baturo, Alexander, Niheer Dasandi, and Slava Mikhaylov. "Understanding State Preferences
with Text as Data: Introducing the UN General Debate Corpus." *Research & Politics*, vol. 4,
no. 2, 2017, pp. 1–9.

Charteris-Black, Jonathan. *Politicians and Rhetoric: The Persuasive Power of Metaphor*.
Palgrave Macmillan, 2011.

Grimmer, Justin, and Brandon M. Stewart. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis*, vol. 21, no. 3, 2013, pp. 267–97.

Mikhaylov, Slava, Alexander Baturo, and Alvaro Dasandi. "United Nations General Debate Corpus: A Resource for International Relations and Comparative Politics." *International Studies Review*, vol. 22, no. 4, 2020, pp. 152–70.

Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. "Thumbs Up? Sentiment Classification Using Machine Learning Techniques." *Proceedings of the ACL Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2002, pp. 79–86.

Roberts, Margaret E., et al. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science*, vol. 58, no. 4, 2014, pp. 1064–82.

United Nations General Debate Corpus 1946–2024. Harvard Dataverse, 2024.

Van Dijk, Teun A. *Discourse and Context: A Sociocognitive Approach*. Cambridge UP, 2008.

World Bank. *World Development Indicators*. World Bank Group, 2024.

Young, Lori, and Stuart Soroka. "Affective News: The Automated Coding of Sentiment in Political Texts." *Political Communication*, vol. 29, no. 2, 2012, pp. 205–31.