INTERNATIONAL MONETARY FUND

From Text to Quantified Insights: A Large-Scale LLM Analysis of Central Bank Communication

Thiago Christiano Silva, Kei Moriya, and Romain Michel Veyrune

WP/25/109

IMF Working Papers **describe research in progress by the author(s) and are published to elicit comments and to encourage debate.**
The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

2025
JUN

WORKING PAPER

© 2025 International Monetary Fund WP/25/109

IMF Working Paper
Monetary and Capital Markets Department

From Text to Quantified Insights: A Large-Scale LLM Analysis of Central Bank Communication
Prepared by Thiago Christiano Silva, Kei Moriya, and Romain Michel Veyrune

Authorized for distribution by Romain Michel Veyrune
June 2025

***IMF Working Papers* describe research in progress by the author(s) and are published to elicit comments and to encourage debate.** The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

ABSTRACT: This paper introduces a classification framework to analyze central bank communications across four dimensions: topic, communication stance, sentiment, and audience. Using a fine-tuned large language model trained on central bank documents, we classify individual sentences to transform policy language into systematic and quantifiable metrics on how central banks convey information to diverse stakeholders. Applied to a multilingual dataset of 74,882 documents from 169 central banks spanning 1884 to 2025, this study delivers the most comprehensive empirical analysis of central bank communication to date. Monetary policy communication changes significantly with inflation targeting, as backward-looking exchange rate discussions give way to forward-looking statements on inflation, interest rates, and economic conditions. We develop a directional communication index that captures signals about future policy rate changes and unconventional measures, including forward guidance and balance sheet operations. This unified signal helps explain future movements in market rates. While tailoring messages to audiences is often asserted, we offer the first systematic quantification of this practice. Audience-specific risk communication has remained stable for decades, suggesting a structural and deliberate tone. Central banks adopt neutral, fact-based language with financial markets, build confidence with the public, and highlight risks to governments. During crises, however, this pattern shifts remarkably: confidence-building rises in communication to the financial sector and government, while risk signaling increases for other audiences. Forward-looking risk communication also predicts future market volatility, demonstrating that central bank language plays a dual role across monetary and financial stability channels. Together, these findings provide novel evidence that communication is an active policy tool for steering expectations and shaping economic and financial conditions.

JEL Classification Numbers:	E52, E58, E43, C55, D83.
Keywords:	Central bank communication; large language models; forward guidance; monetary policy; sentiment analysis
Authors’ E-Mail Addresses:	TSilva@imf.org; KMoriya@imf.org; and RVeyrune@imf.org

1. Introduction

Central bank communication is fundamental to modern monetary policy, playing an important role in shaping market expectations, influencing economic decisions, and enhancing central bank accountability and transparency (Blinder et al., 2008; Woodford, 2005). Effective communication helps central banks' policies to be understood and anticipated by market participants (Bholat et al., 2015; Haldane & McMahon, 2018). In recent years, central bank communication has expanded beyond monetary policy to cover many emerging topics, reflecting growing responsibilities and heightened public scrutiny. In response, many institutions are seeking to modernize their communication strategies to improve policy transmission, strengthen institutional credibility, and safeguard independence by fostering public trust. However, transforming complex language into actionable insights remains challenging, given the intricacies of policy discourse and the public's unfamiliarity with central banking. Systematic and quantitative approaches are increasingly essential in this effort. They would enable central banks to assess the clarity and consistency of their communication, align messages with policy objectives, benchmark performance against peers, and respond more effectively to public concerns—all of which ultimately reinforce trust and institutional legitimacy.[1]

This paper develops an automated classification tool that systematically analyzes central bank communications along four key dimensions–topic, communication stance, audience, and sentiment–offering a comprehensive framework for evaluating policy messages. The topic classification identifies key themes, such as monetary policy, financial stability, and climate change, allowing policymakers to track their messaging focus over time. The communication stance captures whether statements are forward- or backward-looking, an essential factor for managing expectations. The audience classification ensures that communication is assessed in terms of its intended recipients, distinguishing messages directed at financial markets, businesses, households, governments, and international stakeholders. Finally, policy sentiment measures the tone of central bank statements—categorizing them as hawkish, dovish, neutral, risk-highlighting, or confidence-building—thereby offering insights into how policymakers convey their policy intentions. Our classification operates at the sentence level rather than the document level, providing greater flexibility to capture content shifts within the same document and offering a fine-grained analysis of central bank communication.

By integrating four dimensions—topic, communication stance, audience, and policy sentiment—into a unified analytical framework, this paper introduces a systematic and semantically rich characterization of central bank communication. Prior research has made important advances by analyzing specific aspects of communication, such as policy tone through sentiment analysis (Apel et al., 2022; Hansen et al., 2017) or shifts in policy focus via topic modeling (Cieslak & Schrimpf, 2019; Correa et al., 2020). However, these methods typically address each dimension in isolation and are often based on dictionary approaches (Aruoba & Drechsel, 2024; Ehrmann & Talmi, 2020; Shapiro et al., 2022), which struggle to capture contextual nuance. For instance,

[1] The IMF is actively supporting these efforts through technical assistance missions and Central Bank Transparency Code assessments, where tools like the classifier developed in this paper are used to generate quantitative diagnostics, benchmark communication practices against peer countries, and complement experts' knowledge with tailored advice to strengthen transparency and accountability.

the term “tightening” may refer to future monetary policy actions or macroprudential regulation, depending on the surrounding text—distinctions that such methods cannot resolve. Our framework addresses this limitation by leveraging large language models (LLMs), which enable sentence-level understanding of meaning and intent across multiple communication dimensions.

Recent work has already begun to apply LLMs and embedding-based methods to central bank communication. de Araujo et al. (2025) construct word embeddings to study semantic shifts in ECB statements, while Pfeifer & Marohl (2023) and Gambacorta et al. (2024) finetune LLMs to classify monetary policy documents, and Hansen & Kazinnik (2024) explore the use of ChatGPT to interpret policy language. These studies mark a shift toward more semantically informed analyses, but most remain focused on narrow tasks, such as tone or topic classification, or are limited to single-country datasets. In contrast, our framework jointly classifies multiple dimensions within a single architecture trained on a large, multilingual corpus from most central banks worldwide. This design allows for a more holistic and context-sensitive communication interpretation, supporting comparative analyses of messaging strategies across countries and time.

Moreover, while existing textual indicators—such as readability scores and syntactic complexity metrics—offer insights into the form of communication, they fall short of capturing the economic intent embedded in policy language. Effective communication is not solely about *pro-forma* considerations but also about how policymakers convey economic assessments, signal future actions, and justify decisions. Our sentence-level LLM classifier directly addresses this gap by extracting meaning from full statements rather than isolated terms, enabling a richer analysis of how central banks justify decisions and shape narratives. By extending beyond surface features to capture the underlying purpose of communication, this framework provides a scalable tool for systematically evaluating central bank messaging and its evolution in response to shifting economic and institutional contexts.

An innovation of our approach is the direct textual measurement of forward-looking communication in central bank statements, capturing forward guidance and other prospective policy signals explicitly from language rather than inferred indirectly. Unlike conventional approaches—which typically estimate forward guidance as residuals from structural models (Campbell et al., 2012; Nakamura & Steinsson, 2018) or through market reactions around policy announcements (Swanson, 2021)—our sentence-level classification distinguishes prospectively oriented statements from backward-looking assessments in monetary policy communications. Importantly, this semantic measure identifies forward-looking policy signals beyond conventional interest rate guidance, encompassing explicit references to unconventional monetary policy tools, such as asset purchase programs (quantitative easing or tightening), liquidity operations, and balance sheet strategies. Furthermore, our multilingual framework generalizes this analysis beyond major central banks like the United States Federal Reserve (Fed)—traditionally the focus in prior literature (Bundick & Smith, 2020; Hubert & Labondance, 2021)—enabling a comprehensive assessment of how diverse institutions communicate expectations across varied economic and institutional contexts.

Our framework further extends traditional sentiment analysis in central bank communication (Correa et al., 2020; Gorodnichenko et al., 2024; Shapiro et al., 2022), augmenting the conventional hawkish-dovish-neutral (positive-negative-neutral) paradigm with two novel categories: risk-highlighting and confidence-building. This expansion enables a more precise identification of messages that assess economic risks or reinforce stability. Without these distinctions, risk-related statements could be misclassified as negative, while confidence-building messages might be mistaken for positive sentiment, leading to potential misinterpretations of policy intent. For example, a warning about downside risks, though seemingly negative, may serve as a proactive measure to guide market expectations rather than signal a monetary policy stance. Similarly, a reassurance of financial stability during economic distress could be misconstrued as neutral or positive, despite its strategic intent. As financial stability, macroprudential policy, and global interconnectedness gain prominence in central banking, these additional sentiment categories provide a more accurate interpretation of communication.

Empirically, this paper contributes to the literature on central bank communication by constructing, compiling, and processing a comprehensive dataset encompassing 74,882 documents and approximately 21 million sentences from 169 central banks. We invested a significant amount of effort in assembling the dataset and collecting documents from numerous central bank websites. While prior studies often rely on limited datasets focused on single institutions, such as the Federal Reserve or the European Central Bank – ECB (Ehrmann & Fratzscher, 2007; Hansen et al., 2019; Romer & Romer, 2004), or smaller cross-country samples with a limited time frame (Cieslak & Schrimpf, 2019), our dataset spans the period from 1884 to 2025 and includes communication outlets such as monetary policy reports, financial stability reports, annual reports, and press releases for most economies in the world. We also append to this corpus Campiglio et al. (2025)'s consolidated dataset with speeches from 131 central banks from 1986 to 2023, the largest dataset on central bank speeches to date. While most documents are in English, our dataset includes content in multiple languages, reflecting the linguistic diversity of global central banking and enabling a more inclusive analysis of communication practices. This multilingual scope ensures that the study captures regional nuances and provides insights into central banks' localized approaches to global and domestic economic challenges, making it more relevant for previous eras when many central banks did not communicate in English.

We start with a sentence transformer LLM of a moderate size. We opt to fine-tune a general-purpose open-source language model to excel in the domain of central bank communication, rather than using proprietary LLMs with several billion parameters, handcrafted prompt engineering, and text generation tasks. We select an encoder-only sentence transformer to extract rich, dense-vector context-aware representations from sentences. These embeddings encode semantic meaning within the context of central banking, enabling high performance of downstream classification tasks. Using a multilingual sentence transformer further distinguishes our work, enabling analysis across more than 100 languages, including those used by central banks that publish exclusively in their local languages. This multilingual capability enhances the inclusiveness and robustness of our findings, offering a global perspective on central bank communication that is typically overlooked in the related literature.

Our empirical analysis uncovers several key insights. Although monetary policy has consistently been at the core of central bank communications for over a century, primarily due to the widespread mandate for price stability, we observe significant variation in topic emphasis across economies at different stages of development. Advanced economies emphasize financial stability more, while emerging economies focus more on fiscal policy. We observe structural changes in monetary policy communication following the adoption of inflation targeting in many economies.

Backward-looking discussions on exchange rates give room to more forward-looking statements on inflation, interest rates, and economic activity, reinforcing the role of expectation management and output gap assessment in central bank communication. While the financial sector is the primary recipient of central bank communication, we find an increased engagement with international stakeholders and businesses. This finding may reflect both the growing interconnectedness of economies and the need for inflation-targeting regimes to shape expectations across a broader spectrum of recipients. Our sentiment analysis reveals asymmetries in communication, with accommodative policies often accompanied by extensive dovish messaging, whereas restrictive policies are conveyed through more concise, hawkish statements.

We propose four metrics to evaluate central bank communication systematically: the *net policy sentiment*, the *straightforwardness index*, the *explanation index*, and the *net confidence index*. Leveraging a sentence-level classification approach, we decompose each metric into forward- and backward-looking components, offering a nuanced perspective on how central banks signal prevailing conditions versus future policy intentions.

The *net policy sentiment* captures the balance between hawkish and dovish signals, offering a quantitative measure of the directional tone of monetary policy communication. By decomposing this measure into forward- and backward-looking components at the sentence level—a methodological innovation relative to prior aggregate or document-based approaches[2]—we provide new insights into how central banks manage expectations. The forward-looking component reflects not only anticipated policy rate decisions but also unconventional instruments, such as forward guidance and balance sheet policies. Empirically, the forward-looking sentiment systematically predicts future policy rate adjustments and market-based interest rates, confirming its relevance as a key monetary tool. The effects are particularly pronounced for longer-term OIS contracts, where monetary policy expectations are key. By contrast, backward-looking sentiment primarily rationalizes past and current conditions, correlating with contemporaneous policy rates but showing limited association with forward-looking market variables.

Importantly, our cross-country analysis reveals systematic heterogeneity in monetary policy communication strategies. Advanced economies make intensive use of the forward-looking component of the net policy sentiment, especially during crises when conventional tools are constrained and communication plays a critical role in stabilizing expectations. Emerging and low-income economies, in turn, rely more heavily on backward-looking narratives shaped by prevailing conditions and structural limitations. Together, these complement the literature on central bank communication by establishing empirically the relevance of monetary policy communication in shaping market-driven indicators.

The *straightforwardness index* captures the extent to which monetary policy communication conveys unidirectional stance signals, distinguishing direct guidance from more conditional or hedged language. While prior research highlights the importance of clarity for anchoring

[2] Existing empirical studies typically classify entire documents or paragraphs into coarse sentiment categories (e.g., hawkish, dovish, neutral), limiting their ability to distinguish between narratives about past conditions and forward-looking policy guidance. Our approach leverages sentence-level classification to isolate directional policy signals embedded in distinct temporal references, improving granularity and interpretability. See Gorodnichenko et al. (2024), Correa et al. (2020), and Shapiro et al. (2022) for examples of document-level approaches.

expectations and enhancing transmission (Blinder et al., 2008; Coenen et al., 2017), empirical studies have largely overlooked how central banks strategically adjust clarity depending on macroeconomic conditions and institutional environments. Existing measures of communication clarity typically rely on readability or linguistic complexity metrics, which do not differentiate between unequivocal guidance and scenario-based communication. The straightforwardness index instead directly captures this distinction, offering a sharper lens into signaling strategies. Across countries, straightforwardness declines sharply during systemic stress episodes—including the global financial crisis and the COVID-19 pandemic—reflecting a deliberate shift toward conditional and scenario-based communication when uncertainty is high. This pattern supports the theoretical view that preserving flexibility becomes essential under heightened macroeconomic uncertainty, as issuing overly deterministic signals risks damaging credibility should conditions evolve unexpectedly (Campbell et al., 2012; Gertler & Karadi, 2015).

Outside crises, cross-country patterns reveal that straightforwardness varies systematically with institutional and monetary frameworks. Advanced and inflation-targeting economies exhibit lower straightforwardness, consistent with strategic efforts to communicate risk scenarios and maintain flexibility as part of their credibility-building approach. By contrast, emerging and low-income economies favor more explicit and direct statements to stabilize expectations in environments marked by weaker nominal anchors and higher external vulnerabilities. Furthermore, forward-looking communication is consistently less straightforward than backward-looking statements, reflecting the inherent uncertainty in signaling future policy intentions. Together, these results demonstrate that straightforwardness is not a fixed feature of central bank communication but a policy variable actively managed in response to macroeconomic conditions and institutional constraints. Our results complement the literature on central bank communication by empirical evidence that straightforwardness is systematically tailored to balance commitment and flexibility in monetary policy signaling.

The *explanation index* quantifies how central banks justify and contextualize policy decisions, capturing the narrative elaboration that accompanies different phases of the policy cycle. Our results uncover a systematic asymmetry: explanation rises sharply during tightening and normalization phases—especially in advanced economies—when restrictive measures require stronger justification to reinforce credibility and anchor expectations, while it declines during easing cycles, such as the onset of the COVID-19 pandemic, when accommodative actions are more readily accepted. Although explanation levels vary across countries, being higher on average in low-income and pegged exchange rate economies, likely due to weaker institutional credibility, these differences narrow during global monetary cycles, underscoring the responsiveness of explanatory communication to systemic shocks rather than structural factors.

The *net confidence index* captures the balance between confidence-building and risk-highlighting language, offering distinctive insights when decomposed into forward- and backward-looking components. Backward-looking confidence reflects assessments of current and past macro-financial conditions, corroborated empirically by the fact that implicit market volatility (VIX) predicts subsequent shifts in this component. Forward-looking confidence, in turn, captures central banks' views about future risks and serves as an active channel for shaping expectations: it predicts future market volatility, demonstrating that risk communication is not merely descriptive but an integral tool of policy signaling. This perspective is related to the approach of Cieslak et al. (2023), who analyze how policymakers' perceived uncertainty—quantified from FOMC transcripts—affects monetary decisions after controlling for the hawkish-dovish tone and economic projections. While both approaches emphasize the importance of uncertainty in shaping monetary policy, our framework differs in scope and design: (i) it systematically quantifies risk-related communication across 169 central banks; and (ii) it introduces a forward/backward decomposition that enables a more granular understanding of how central banks communicate risk, whether to explain past conditions or to signal future developments.

Finally, we provide systematic empirical evidence on how central banks tailor risk communication across audiences, advancing prior work that largely relied on anecdotal or descriptive assessments. We show that audience differentiation is structural and persistent. Central banks communicate with the general public by building confidence to reinforce trust and anchor expectations, while adopting a more risk-oriented tone with governments to highlight vulnerabilities critical for fiscal prudence. Communications with the business and financial sectors are more balanced, avoiding excessive optimism or pessimism to prevent misinterpretation and destabilizing market reactions. This audience targeting evolves in response to systemic stress. During crises, central banks shift towards building confidence, particularly in messages directed at the financial sector and government, aiming to reassure key actors essential to crisis mitigation and policy transmission. By contrast, communication with the general public and international stakeholders becomes more cautious, emphasizing risks and uncertainties. Despite these cyclical adjustments, the overall pattern of differentiated tone across audiences remains remarkably stable over decades, underscoring that targeted sentiment is not merely reactive but a deliberate and enduring feature of central bank communication strategy.

The paper is structured as follows. Section 2 details the sample selection criteria, dataset compilation, and preprocessing procedures. Section 3 presents an analysis of the central bank communication dataset, focusing on textual form measures, including readability and syntactic complexity. In Section 4, we shift to a semantic analysis using a fine-tuned classifier, categorizing central bank statements by topic, communication stance, audience, and sentiment. Section 5 defines semantic textual metrics–net policy sentiment, straightforwardness index, explanation index, and net confidence index–and shows their application to the dataset. Finally, Section 6 offers concluding remarks and highlights potential directions for future research.

2. The Central Bank Communication Dataset

2.1. Data Collection

We classify documents into two types: regular and non-regular central bank communication outlets. Annual reports, monetary policy decisions (statements, press releases, and minutes), monetary policy reports, and financial stability reports comprise the regular suite of documents typically produced by central banks and often required by domestic legislation. Other documents include speeches, specialized reports, and press releases. Table 1 lists the collected documents, their institutional purposes, and the range of data availability.

We compiled published documents from 169 central bank websites, ensuring a broad representation of central bank communications. While most of these documents are in English, the dataset also contains documents in several other languages, reflecting the global scope of central

Table 1: Types of Central Bank Communication Outlets Used in the Analysis

Category	Document Type	Description	Number Docs	Begin Year	End Year
Regular	Annual Report	Key institutional communication detailing central bank governance, financial statements, economic developments, and policy implementation. Often required by legislation.	3,879	1884	2024
Regular	Monetary Policy Report[2]	Overview of the central bank's monetary policy stance, actions, and economic outlook. Typically published quarterly.	4,671	1993	2025
Regular	Financial Stability Report[2]	Semiannual or annual report evaluating financial sector risks and vulnerabilities.	2,092	1996	2025
Regular	Monetary Policy Decision[3]	Communication issued after interest rate decisions detailing the rationale behind the policy move.	14,238	1936	2025
Non-Regular	Speeches	Speeches by central bank decision-makers, often on economic and monetary policy matters. The majority of the data comes from Campiglio et al. (2025).	36,725	1986	2025
Non-Regular	Other Documents[4]	Press releases and reports on specialized topics that do not fall into the above categories.	11,437	1993	2025

[1.] Monetary policy reports are called inflation reports in some economies. A few specific countries publish both.
[2.] Financial stability reports are called financial stability reviews or financial stability surveillance in some economies.
[3.] Monetary Policy Decisions include press releases, official statements, and meeting minutes:

- *Press Release:* Announced after each monetary policy decision, providing a summary of the decision and rationale.
- *Statement:* Official statement outlining the central bank's position, economic analysis, and expectations.
- *Minutes:* Detailed records of the discussions held during policy meetings, offering insights into decision-making processes.

[4.] Other Documents encompass various central bank specialized reports and press releases, including bulletins, balance of payment reports, economic and monetary reports, and specialized publications on monthly reviews, economic outlook, monetary policy data, macroeconomic reviews, market operations and monetary policy.

bank communications. We only considered official translations into English that the central bank explicitly published. Otherwise, the document is processed in its local language. Most of the speeches in our data set originate from the very comprehensive compilation by Campiglio et al. (2025), who collected 35,487 unique speeches from 131 central banks from 1986 to 2023, the largest dataset on central bank speeches to date. We augment this comprehensive dataset with speeches published January 2024 onward using the consolidated dataset maintained by the BIS.

We ingested the collected documents into a comprehensive pre-processing pipeline designed to extract, segment, and standardize text from a diverse range of file formats, including PDFs, Word documents (.docx, .doc), plain text files (.txt), and HTML files. Image-based documents, such as scanned PDFs and embedded images within Word files, were processed using optical character recognition (OCR) to recover textual content. HTML documents require additional processing to extract only the core textual content from entire web pages, removing navigation menus, banners, and other non-relevant elements. We implemented this post-processing to ensure we retain only substantive communications from central banks.

For the classification part, we work at the sentence level rather than the document level. The sentencization process is not straightforward due to language-specific idiosyncrasies. Sentence boundaries vary significantly across languages, requiring tailored segmentation approaches to ensure accuracy. We address this by applying language detection and then employing language-specific segmentation models optimized for different grammatical structures and punctuation conventions. This approach minimized sentence segmentation errors that could arise from ambiguous abbreviations, varying sentence-ending markers, and other linguistic idiosyncrasies. Since transformer-based models are highly effective at capturing semantic meaning from raw text, we applied minimal textual modifications beyond standardization to maintain consistency across sources.

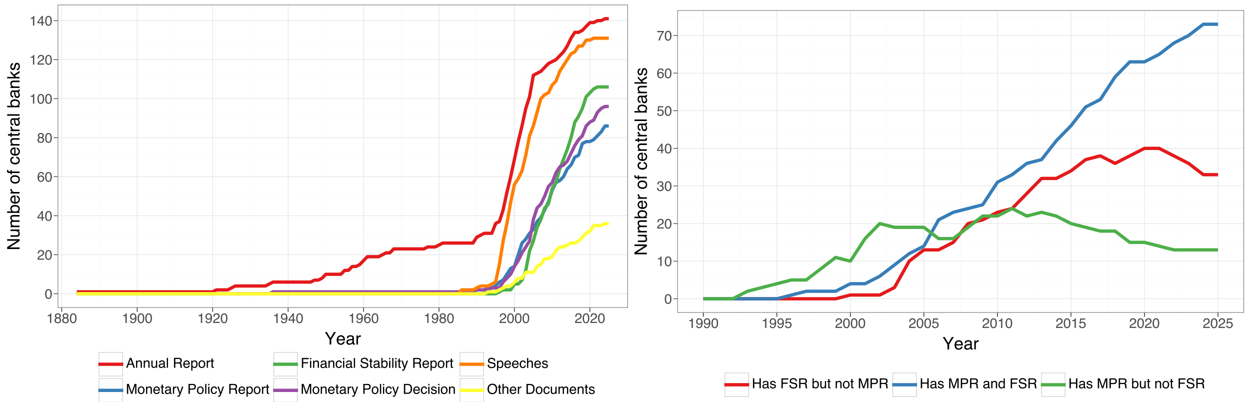
2.2. Data Exploratory Analysis

Our dataset comprises 74,882 central bank documents, of which 24,880 correspond to regular documents of central bank communication. This corpus consists of over 80 GB of textual information, spans more than 1.3 million processed PDF pages (excluding other file formats), and encompasses approximately 21 million processed sentences. The breadth and depth of these publications enable an unprecedented analysis of central bank communication in many dimensions, including monetary policy and financial stability. Figure 1a illustrates the number of central banks that have published central bank communication instruments over time, highlighting the dataset's temporal and cross-country coverage. While English documents dominate (87.4 percent of sentences), the dataset spans over 30 languages, including Spanish (6.6 percent), Portuguese (1.7 percent), French (1.4 percent), Arabic, Russian, Chinese, and others. This linguistic diversity underscores the global coverage of the analysis and the inclusiveness of our approach across distinct communication ecosystems.

The dataset exhibits a significant imbalance over time, with some central banks maintaining a long and consistent record of publications, allowing for the study of evolving communication strategies over decades. Bulgaria holds the most extensive time series of annual reports, dating back to 1884.[3] The United States has the longest recorded time series of monetary policy minutes and policy actions, dating back to 1936, with a relatively high frequency. The United Kingdom provides the most extensive series of financial stability reports, dating back to 1996. The earliest monetary policy reports in the dataset originate from Sweden and the United Kingdom, both of which date back to 1993.

[3] Other countries with extensive annual report time series include Finland (1921), Mexico (1925), Chile (1926), Argentina (1935), India (1936), the United Kingdom (1936), the Philippines (1949), Sri Lanka (1950), Brazil (1955), Israel (1955), Germany (1957), Suriname (1957), Tunisia (1959), Australia (1960), Nigeria (1960), Saudi Arabia (1961), Qatar (1966), Uganda (1967), Mauritius (1968), Belize (1977), Bolivia (1980), and Switzerland (1981).

Figure 1: Trends in Central Bank Publications



(a) Number of central banks publishing in different communication outlets.

(b) Publication trends of MPRs and FSRs.

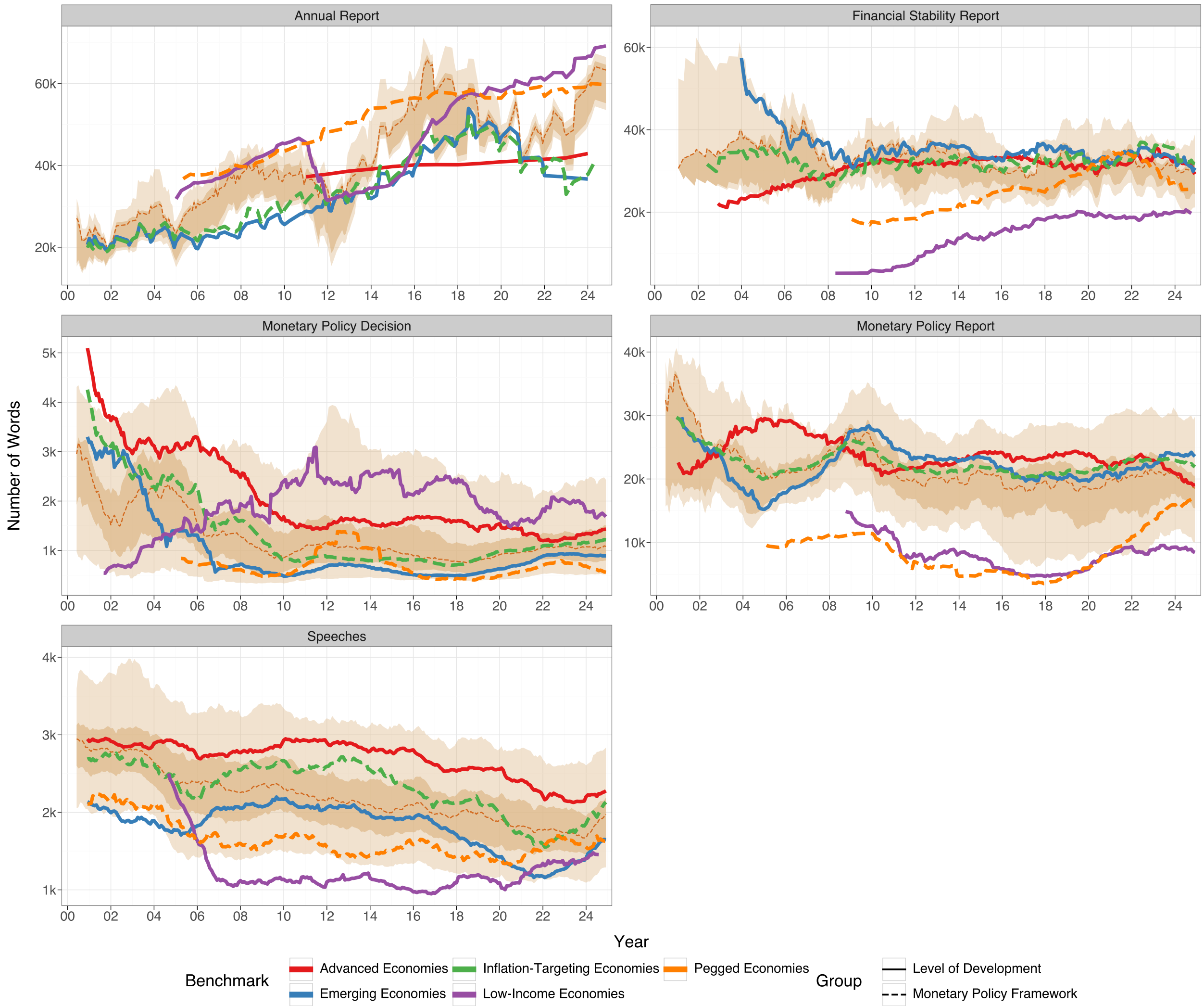
Notes: The left panel shows the number of central banks that have published central bank communication instruments over time, where a central bank is counted at time t if it has published the communication outlet at least once up to time t (resulting in a non-decreasing curve). The right panel displays the number of central banks publishing monetary policy reports (MPRs) and/or financial stability reports (FSRs) over time, highlighting differences in adoption.

Figure 1b shows that an increasing number of central banks are publishing monetary policy reports and financial stability reports. Monetary policy reports became more prevalent as the inflation-targeting monetary policy framework gained popularity. The global financial crisis also prompted central banks to prioritize financial stability. Hence, the number of central banks that publish financial stability reports jumped after the global financial crisis. Interestingly, financial stability reports are more prevalent than monetary policy reports, particularly in economies with exchange rate anchor regimes—such as Brunei Darussalam, Nepal, and Singapore—where monetary policy decisions are limited compared to other regimes, reducing the need for dedicated monetary policy reports.

Figure 2 shows the evolution in the number of words across five key central bank communication outlets—annual reports, financial stability reports, monetary policy reports, monetary policy decisions, and speeches—disaggregated by level of development and monetary policy framework. The global interquartile range (25^{th}–75^{th} percentile, shaded area) serves as a benchmark for comparative positioning. Annual reports have grown significantly in length over time, particularly among low-income and pegged economies. This reflects both the expansion of central bank functions in these jurisdictions and the reliance on annual reports as their primary or sole outlet for communicating monetary, financial, and institutional developments. In contrast, advanced, emerging, and inflation-targeting economies have streamlined their annual reports since 2018, consistent with their more differentiated publication strategies and more mature institutional communication practices. Financial stability and monetary policy reports show similar dynamics: longer formats in advanced and inflation-targeting economies, and shorter but gradually expanding ones in others, reflecting both capacity constraints and varying degrees of institutional evolution.

Monetary policy decisions were being streamlined in advanced and inflation-targeting economies from 2000 until the COVID-19 pandemic. However, this trend reversed after the global inflation surge, as elevated uncertainty and the constrained policy space of traditional policy tools required more detailed narrative explanations to re-anchor expectations and justify policy actions. Speech lengths followed a similar U-shaped pattern: they declined initially, then increased again as central banks made greater use of this channel. The consistently higher speech length in advanced and inflation-targeting economies likely reflects their more active use of speeches as tools for forward guidance, market calibration, and stakeholder engagement.

Figure 2: Document Length Trends by Development Level and Monetary Policy Framework



Notes: This figure shows the average number of words over time in five types of central bank documents—annual reports, financial stability reports, monetary policy reports, monetary policy decisions, and speeches. Results are disaggregated by level of development (solid lines) and monetary policy framework (dashed lines). The shaded area shows the interquartile range (25^{th}–75^{th} percentile) of the global distribution; the dashed brown curve represents the global median. Averages are calculated across all documents published by the same central bank in each outlet.

3. *Pro-Forma* Analysis of Central Bank Communications

This section examines the *pro-forma* characteristics of central bank communication, focusing on lexical readability metrics (e.g., Flesch-Kincaid scores) and structural complexity indicators of sentences (e.g., syntactic dependency depth). These measures provide valuable insights into the surface accessibility of policy messages—how easily they can be read and syntactically parsed by target audiences. However, they do not capture the semantic content of communication, including the underlying economic rationale, policy intent, or strategic framing. While readability and syntactic structure shape the cognitive effort required to process information, effective communication also depends critically on *what* is communicated and *how* policy narratives are constructed. The following section will address this aspect by leveraging large language models to analyze the semantic dimensions of central bank communication, enabling a more in-depth assessment of policy communication content.

Figure 3 portrays the lexical readability by communication outlet using the Flesch-Kincaid Ease score disaggregated by level of development and monetary policy framework for English documents.[4] The Flesch Reading Ease metric assigns a readability score for the text based on two components: the sentence length and the word complexity (measured in terms of the average syllables per word). The higher these components, the lower the readability. The figure also illustrates these two components, enabling us to understand the factors that drive changes in lexical readability over time.

Documents from pegged exchange rate regimes and low-income economies consistently display higher Flesch Reading Ease scores, driven by shorter sentences and simpler word choices. This finding is consistent with the limited discretion and policy complexity of these regimes, where communication typically focuses on operational updates rather than forward-looking guidance. In contrast, inflation-targeting and advanced economies consistently score lower in readability, particularly in reports on monetary policy and financial stability. These documents are characterized by longer sentences and more complex vocabulary, reflecting the technical nature of communication required to explain forward-looking strategies, manage expectations, and signal credibility under discretionary regimes. Notably, the widening readability gap across development levels suggests

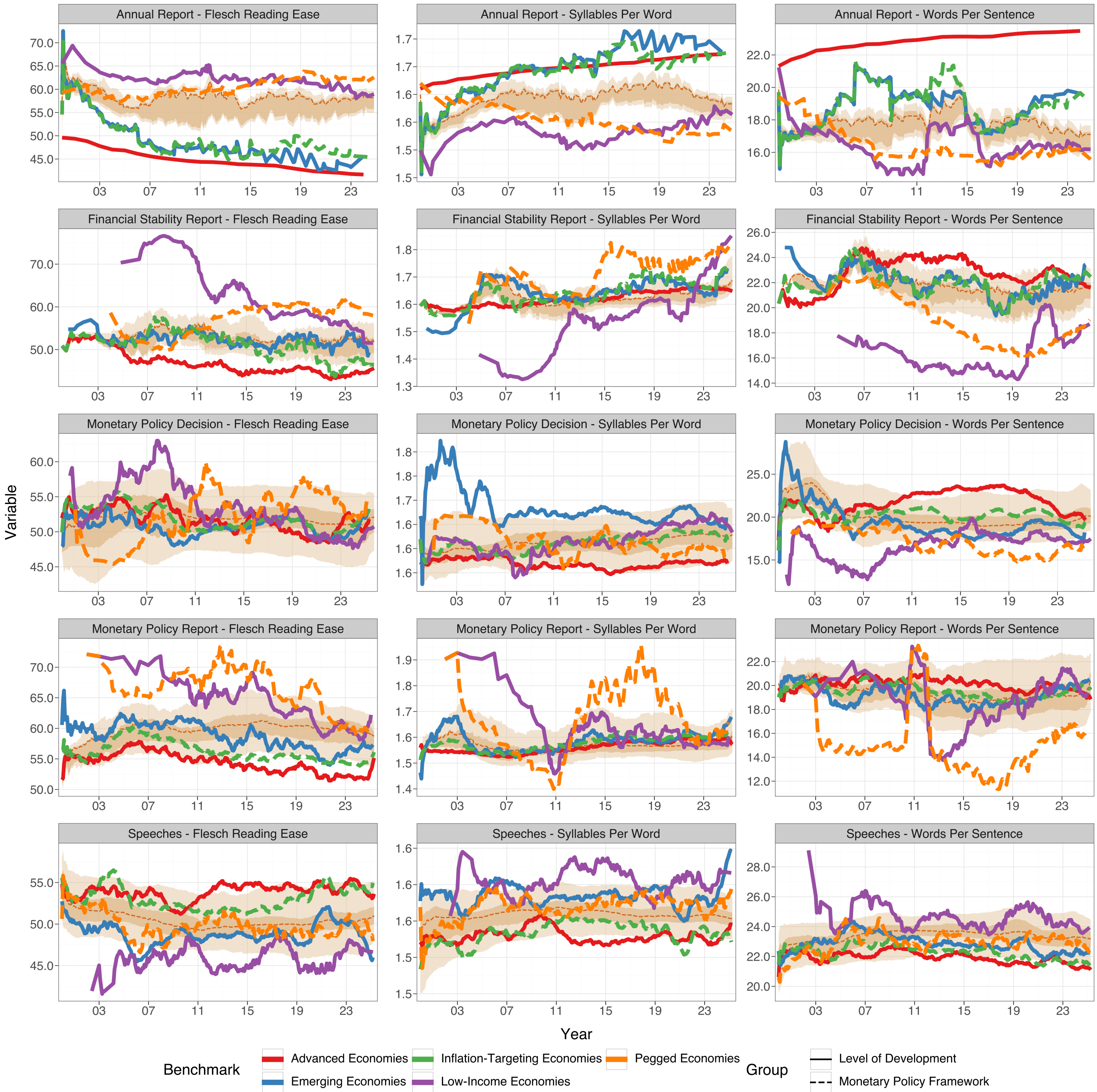
[4] The Flesch-Kincaid Ease Score is calculated as:

$$\text{Score} = 206.835 - 1.015\left(\frac{\text{Total Words}}{\text{Total Sentences}}\right) - 84.6\left(\frac{\text{Total Syllables}}{\text{Total Words}}\right)$$

The score ranges from 0 to 100, with higher values indicating better readability. In English-language texts, scores above 60 correspond to an 8th-grade reading level, while values between 30 and 50 suggest college-level complexity. Scores below 30 indicate highly technical or specialized content. However, direct comparability is limited since our dataset includes documents in multiple languages. Different languages exhibit structural variations—such as average word length and syntactic complexity—that affect readability scores differently. For instance, agglutinative languages (e.g., Finnish, Turkish) and languages with complex morphology (e.g., German) tend to yield lower readability scores than more analytically structured languages, such as English or Chinese. Despite these variations, the metric remains useful for assessing relative trends within each language and across central bank communication types. While it is possible to calculate this metric in different languages, for comparison purposes, only English documents were used in this part of the analysis.

that as central bank mandates grow more sophisticated, so too does the complexity of their public communication, raising important considerations for accessibility and stakeholder engagement.

Figure 3: Lexical Readability Trends by Development Level and Monetary Policy Framework



Notes: This figure tracks the evolution of lexical readability (left), syllables per word (center), and words per sentence (right) for five types of central bank communication: annual reports, financial stability reports, monetary policy decisions, monetary policy reports, and speeches. Readability is proxied by the Flesch-Kincaid Ease score, where higher values indicate simpler language. For the component metrics, higher values denote greater linguistic complexity. Results are disaggregated by level of development (solid lines) and monetary policy framework (dashed lines). Shaded areas show the global interquartile range (25^{th}–75^{th} percentile); the dashed brown line represents the global mean.

We also analyze the sentence structure of published documents, focusing on the degree of syntactical complexity. Specifically, we examine the dependency depth of sentence structures, which captures the extent to which words in a sentence are hierarchically nested within syntactic trees. Greater depth indicates more complex sentence constructions, often characterized by embedded clauses and long-distance dependencies, which can hinder readability and comprehension. For instance, the simple sentence "Central banks adjust interest rates" has a shallow dependency tree with a depth of 2, whereas the more complex sentence "Central banks, in response to inflationary pressures, swiftly adjust interest rates to maintain stability" exhibits a depth of 4 due to the additional nested phrases.[5] Unlike lexical readability metrics, dependency depth accounts for sentence structure beyond word-level properties, making it particularly relevant for assessing clarity in central bank communication.

Figure 4 presents average sentence structure complexity—measured by syntactic dependency depth—across five types of central bank communications, disaggregated by level of development and monetary policy framework. A key insight emerges when these patterns are contrasted with lexical readability metrics. While advanced economies consistently use more complex vocabulary (lower lexical readability), they tend to structure their messages with simpler sentence constructions. This is particularly evident in annual reports, financial stability reports, and monetary policy reports, where these economies maintain consistently lower syntactic dependency depth. The deliberate choice to convey technically rich content through syntactically simple sentences likely reflects an effort to balance transparency with accessibility, preserving precision while facilitating comprehension among broader audiences.

In contrast, low-income and pegged economies often display the inverse pattern. Their documents exhibit significantly deeper sentence structures despite addressing less technically complex issues. This suggests a more convoluted exposition style that may compromise clarity of communication. Most strikingly, the gap in sentence structure complexity has widened in monetary policy decisions, the most prominent outlet for monetary policy communication. Since 2020, advanced economies have continued to simplify their sentence structure amid post-pandemic uncertainty. In contrast, low-income economies have moved in the opposite direction. The sharp increase in sentence complexity beginning in 2022 coincides with the global surge in inflation, suggesting that these economies may face heightened difficulties in communication during macroeconomic stress periods. While advanced economies appear able to adapt by making

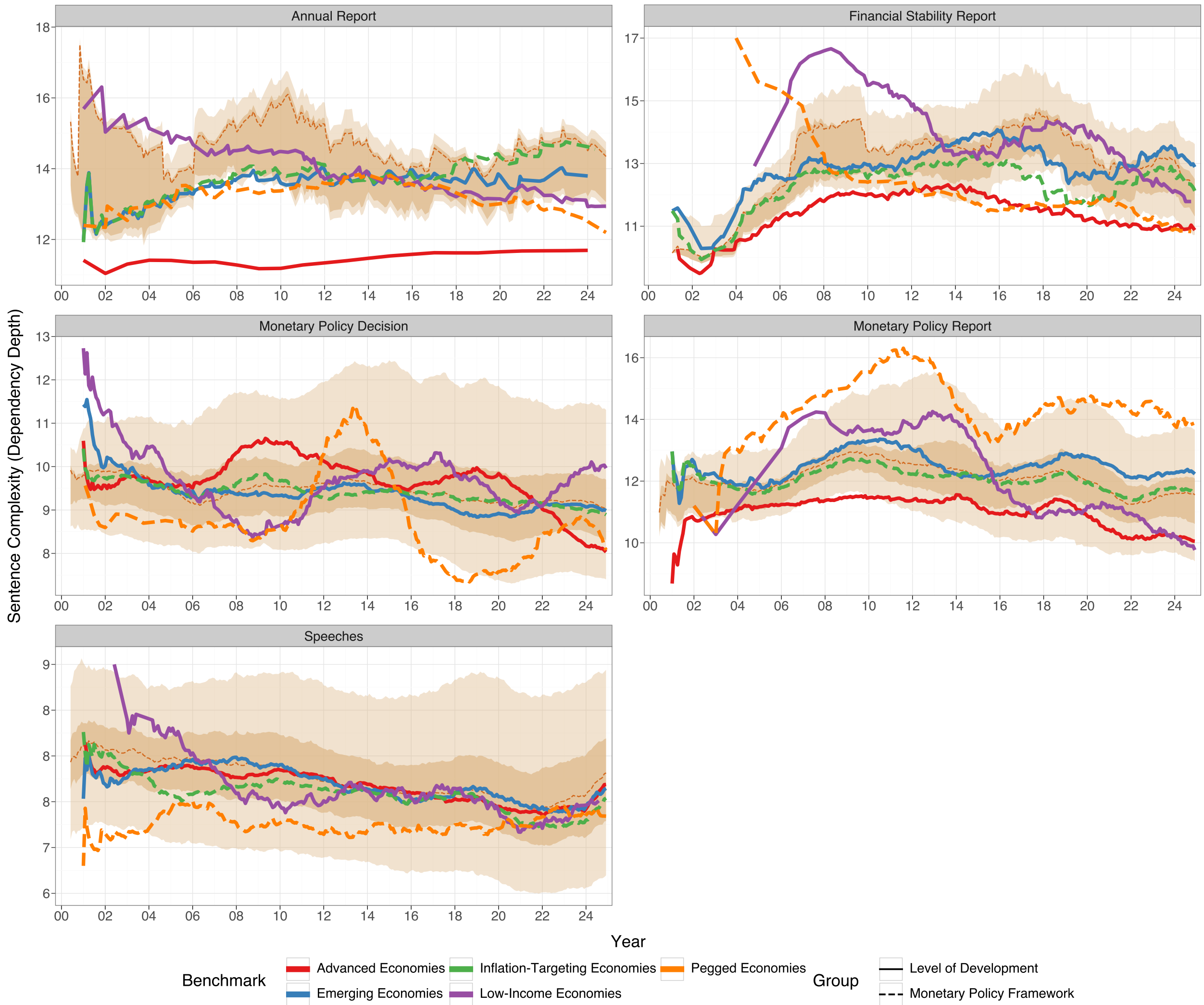
[5] Dependency depth is measured as the longest path from the root of a sentence's syntactic tree to any of its terminal nodes. Formally, given a dependency tree T with root r and a set of leaf nodes L, the depth of a sentence is computed as:

$$\text{Depth}(T) = \max_{l \in L} \text{distance}(r, l).$$

For example, consider the sentence "The central bank raised interest rates." The syntactic tree consists of the root verb "raised," with direct dependencies to the subject "bank" (which itself depends on "The") and the object "rates" (which is modified by "interest"). The longest dependency path, in this case, is from "raised" to "The," yielding a depth of 3. By contrast, a more complex sentence such as "The central bank, recognizing inflation risks, decided to raise interest rates preemptively" introduces additional dependency layers, increasing the depth to 5. Higher values indicate greater syntactical complexity, as longer paths indicate deeper layering of phrases and clauses. Cross-linguistic differences must be considered, as languages with freer word order or richer morphology (e.g., German, Russian) naturally exhibit higher syntactic depth than more rigidly structured languages (e.g., English, Chinese).

their communication even more accessible in high-volatility environments, low-income central banks may lack the institutional tools or technical capacity to do so, potentially undermining transparency.

Figure 4: Sentence Structure Complexity by Development Level and Policy Framework

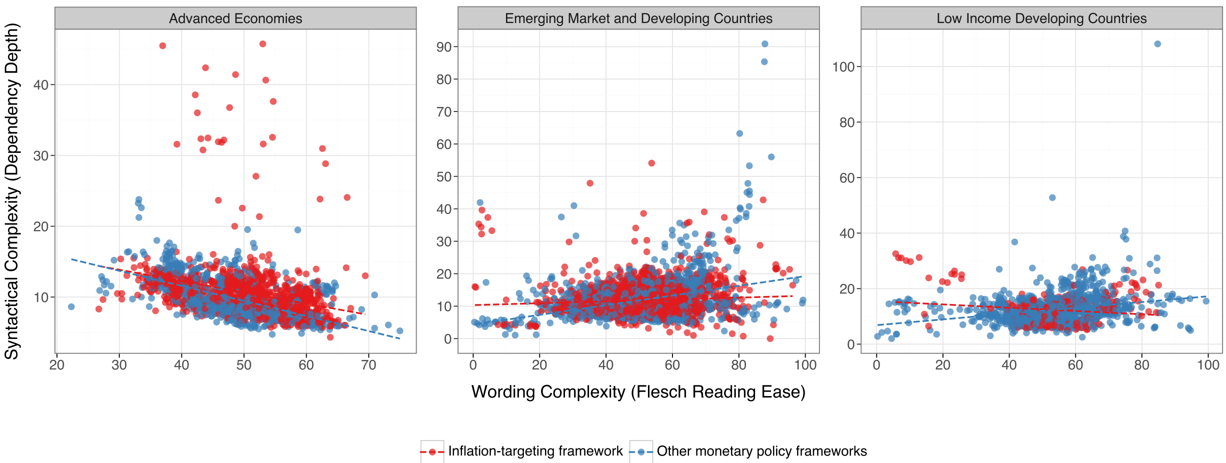


Notes: This figure presents the average sentence structure complexity by communication type, measured using syntactic dependency depth. Each panel corresponds to a specific communication outlet: annual reports, financial stability reports, monetary policy decisions, monetary policy reports, and speeches. Higher values indicate deeper, more nested syntactic constructions. Results are disaggregated by level of development (solid lines) and monetary policy framework (dashed lines). Shaded areas show the global interquartile range (25^{th}–75^{th} percentile) for each year.

These findings underscore that clarity in central bank communication is shaped not only by the complexity of the content but also by institutional capacity to deliver messages in an accessible and well-structured manner. Structural simplicity—especially when paired with technically rich content—is a hallmark of more effective communicators, reinforcing transparency and accountability.

Importantly, lexical and syntactic complexity are not mechanically related. As shown in Figure 5, advanced economies tend to combine more complex vocabulary with clear sentence structure, reflecting a deliberate effort to communicate technical content in a digestible form. In contrast, many low-income and non-inflation-targeting economies rely on simpler vocabulary but construct sentences in more convoluted ways, which may impede clarity and reduce communicative effectiveness.

Figure 5: Lexical and Syntactic Complexity Statistical Relationship



Notes: Each panel in the figure shows a scatter plot with the relationship between lexical and syntactic complexity for countries grouped by development level: advanced economies (left), emerging market and developing countries (center), and low-income developing countries (right). Colors differentiate inflation-targeting economies from other types of monetary policy frameworks. Lexical complexity is measured by the Flesch-Kincaid Ease score, and syntactic complexity is measured by sentence dependency depth. Each dot represents country-specific average values across all communication types for a given time.

While these *pro-forma* indicators offer insights into the surface features of central bank communication, they remain fundamentally limited. They do not capture meaning, intent, or rhetorical strategy—elements that are central to understanding how policy messages are framed and interpreted. The following section introduces a semantic analysis framework based on a large language model that has been fine-tuned for central bank communication to address these limitations. This approach enables a more profound and more systematic assessment of the underlying policy narratives, forward guidance, and institutional messaging strategies.

4. Semantic Analysis of Central Bank Communications

This section outlines the LLM-based central bank communication classification methodology and its application to the large dataset of central bank communications previously discussed.

4.1. Methodology

Here, we outline the methodological steps for selecting, designing, fine-tuning, and validating the large language model for sentence-level central bank classification.

4.1.1. Selection of the LLM

Our methodology entails fine-tuning a general-purpose large language model to classify and analyze central bank communications. Given the specialized nature of central bank discourse—characterized by technical economic language, nuanced policy signals, and varied audience targeting—a domain-specific adaptation is necessary. We opt for a *sentence transformer*—an encoder-only model—rather than a decoder-only autoregressive language model, such as GPT. Sentence transformers generate dense, semantically meaningful embeddings, which are directly optimized for classification and similarity tasks, making them more appropriate for our application than token-by-token generative models like GPT (Reimers & Gurevych, 2019).

Several considerations motivate this design choice. First, sentence transformers utilize bidirectional encoder architectures, enabling full contextualization by simultaneously attending to both the left and right contexts within a sentence (Devlin et al., 2019). This makes them well-suited for sentence-level semantic tasks such as classification, textual entailment, and document similarity. Decoder-only models, such as GPT, are optimized for autoregressive generation and perform best when tasked with open-ended text completion or synthesis. While they can be repurposed for classification through prompt engineering or few-shot learning (Brown et al., 2020), such methods tend to be brittle, require extensive manual calibration, and lack transparency in how label boundaries are inferred—limitations that undermine their suitability for highly structured classification tasks.

Second, sentence transformers are explicitly trained to produce fixed-size embeddings that map semantically similar sentences close together in vector space. These embeddings are learned through contrastive objectives that preserve global semantic relationships (Reimers & Gurevych, 2019). Unlike traditional transformer encoders, which produce token-level outputs that require additional pooling strategies, sentence transformers directly output sentence-level representations. This is particularly critical in central bank communication, where interpretive content relies heavily on sentence-level structure. For example, the sentence "While inflation remains elevated, policy tightening is expected to restore price stability in the medium term" conveys a conditional monetary policy stance that emerges only when considering the sentence as a whole.

Third, from a computational standpoint, encoder-only models offer substantially greater cost and computing efficiency. Inference with GPT models incurs higher costs because pricing is based on the combined length of input prompts and generated output. This cost structure becomes prohibitive for large-scale applications involving millions of sentences, such as our multilingual corpus of central bank documents. Model adaptability over time is also essential in our setting, where taxonomies evolve and new policy themes may emerge. With sentence transformers, updating the model to accommodate new classes or refinements to the label space can be handled through incremental fine-tuning using a relatively small number of curated examples, or even complete fine-tuning. In contrast, adapting GPT models to systematically incorporate new label definitions—especially when precision is critical—would either require full instruction tuning, which is computationally intensive, or prompt re-engineering, which is inherently heuristic and often yields unstable performance across languages and use cases.

Another critical consideration is the multilingual nature of central bank communication. Although major central banks communicate primarily in English, some smaller and emerging-market central banks issue key documents in their native languages. Ignoring multilingualism could introduce selection biases in the analysis, as central bank policy messages can vary in tone, sentiment, or emphasis depending on the language used (Siddhant et al., 2020). This feature is underscored in our dataset, for example, where Mexico and Chile have annual reports dating back to 1925 and 1926, respectively, whereas English reports only became available in 2000.

Given those features, we select and fine-tune a multilingual BGE (bge-m3) sentence transformer explicitly trained to produce high-quality cross-lingual sentence embeddings that align semantically similar statements regardless of language (Chen et al., 2024). This property ensures a consistent classification applied to original local language documents, mitigating potential distortions and subjective biases introduced by translation models, which can alter the economic intent of the original text (Conneau et al., 2020). Appendix B explores this cross-lingual consistency.

4.1.2. Labeled Dataset Construction

We follow a supervised learning approach where we fine-tune the pre-trained LLM using a labeled set with pre-specified labels. Our labeled set comprises 1,200 annotated sentences, constructed through a multi-stage process. The initial draft of sentence examples—comprising approximately 240 instances (20 percent)—was generated using a generative AI model prompted to produce content satisfying two essential properties: (i) semantic differentiation across labels—ensuring that sentences belonging to different topics or classes exhibit distinct semantic features; and (ii) intra-label diversity—ensuring a broad variety of expressions within each class. These synthetic examples provided a scaffold for initial model supervision but were not used without expert verification. See Prompt A.1 for the prompt definition. In the second stage, three domain experts created approximately 240 examples (20 percent) through bespoke sentence construction, informed by practical experience with central bank communication. This set was constructed after reviewing the sentences generated by the chatbot. These expert-generated sentences enriched the dataset with linguistic precision and policy-relevant framing that was not easily captured by automated generation. Finally, the remaining 720 examples (60 percent) were composed of actual sentences extracted from central bank documents. The domain experts carefully selected and manually annotated these to reflect the authentic language and economic reasoning present in real policy documents.

The training set comprises all generated sentences (240), expert-constructed sentences informed by practical experience (240), and half of the sentences extracted from actual central bank communications (360). The remaining 360 real sentences are allocated to a hold-out validation set, ensuring a clean separation between training and evaluation. Constructing the validation set exclusively from real central bank content aligns with best practices in model assessment, as it mirrors the distribution the model will encounter in deployment. If the validation set is not aligned with the target distribution, performance metrics risk being inflated, unstable, and unreliable.

In contrast, the training distribution does not need to precisely mirror the test-time distribution, provided it supplies a sufficiently rich and structured signal to learn transferable representations. This flexibility motivates our hybrid strategy: synthetic and expert-crafted sentences offer fine-grained semantic control and full coverage of the classification taxonomy, while including real sentences supports empirical grounding. This combination enables efficient supervision, especially in domains where naturally labeled examples are limited or imbalanced.

One may question why the training set does not exclusively consist of real sentences from central bank documents. While theoretically appealing, this alternative presents several practical challenges. First, relying solely on real-world text risks introducing sampling bias and coverage gaps, as actual documents may underrepresent rare but economically relevant combinations of topic, communication stance, audience, and sentiment. Second, constructing a fully balanced, labeled dataset from central bank texts alone is prohibitively labor-intensive due to the need for expert annotation. Incorporating generative and expert-created content ensures comprehensive label representation and sentence diversity while reducing annotation costs.

Such an approach is consistent with foundational results in representation learning and domain generalization: performance should be evaluated on the target distribution, while training data may originate from a broader mixture, so long as it induces representations that transfer effectively (Ben-David et al., 2010; Hendrycks et al., 2021; Recht et al., 2019).[6]

4.1.3. Classification Structure

To capture the multidimensional nature of central bank communication in a tractable yet semantically meaningful manner, we train two distinct classifiers: one tasked with jointly predicting topic and communication stance, and another with jointly predicting audience and sentiment. This modeling decision reflects both economic intuition and the empirical structure of the data, allowing us to preserve relevant dependencies while avoiding the sparsity problems inherent in high-dimensional classification.

Formally, let $Y = (y_{\text{topic}}, y_{\text{stance}}, y_{\text{audience}}, y_{\text{sentiment}})$ denote the full set of labels, where y_{topic} corresponds to the topic label, y_{stance} to the communication stance, y_{audience} to the audience, and $y_{\text{sentiment}}$) to the sentiment. Ideally, one would seek to estimate the full joint conditional distribution $P(y_{\text{topic}}, y_{\text{stance}}, y_{\text{audience}}, y_{\text{sentiment}} \mid x)$, where x represents the input sentence embedding. However, doing so would require an impractically large amount of labeled data, as the number of possible label combinations grows exponentially. In our case, estimating this full joint distribution would entail learning hundreds of sparse four-way combinations, many of which are rarely or unobserved in the empirical distribution.

To address this challenge, we adopt a block factorization that assumes conditional independence between two pairs of labels:

[6] Recht et al. (2019) show that even models trained on large-scale datasets such as ImageNet exhibit degraded performance when evaluated on new test sets drawn from the same nominal distribution, underscoring the importance of aligning validation data with the deployment environment. Hendrycks et al. (2021) show empirically that training on a broader or augmented dataset—even one that differs from the test distribution—can enhance generalization, provided that evaluation is conducted on data representative of the deployment setting. Ben-David et al. (2010) provide theoretical guarantees for domain adaptation, showing that the expected error on a target domain can be bounded by the sum of the source error, a divergence measure between the source and target distributions, and the error of the optimal joint hypothesis—i.e., the best hypothesis in the class that performs well on both domains. This bound is meaningful when such a hypothesis exists and the domain divergence is sufficiently small. In our case, the training set combines synthetic, expert-constructed, and real examples to induce semantically transferable representations. In contrast, the validation set consists solely of real central bank communications to ensure alignment with the model's target application.

$$P(y_{\text{topic}}, y_{\text{stance}}, y_{\text{audience}}, y_{\text{sentiment}} \mid x) \approx P(y_{\text{topic}}, y_{\text{stance}} \mid x) \cdot P(y_{\text{audience}}, y_{\text{sentiment}} \mid x). \quad (1)$$

The choice of this factorization—⟨topic, communication stance⟩ and ⟨audience, sentiment⟩—is not arbitrary. To inform this design, we examined the empirical co-occurrence patterns in the labeled dataset, which comprises real sentences from central bank communications. We compared the number of unique label combinations for each pair of label dimensions to a null distribution generated via random permutation, yielding a Z-score that quantifies the degree of structured co-occurrence beyond statistical independence.[7] Figure 6 shows the Z-scores of pairwise empirical coupling for all combinations. The pair *topic* + *guidance* exhibited the highest absolute Z-score, followed by *topic* + *audience*, and then *audience* + *sentiment*. Since the same dimension cannot appear in more than one classifier, we selected the combination of *topic* + *guidance* and *audience* + *sentiment* as the disjoint pair with the highest combined absolute Z-score.

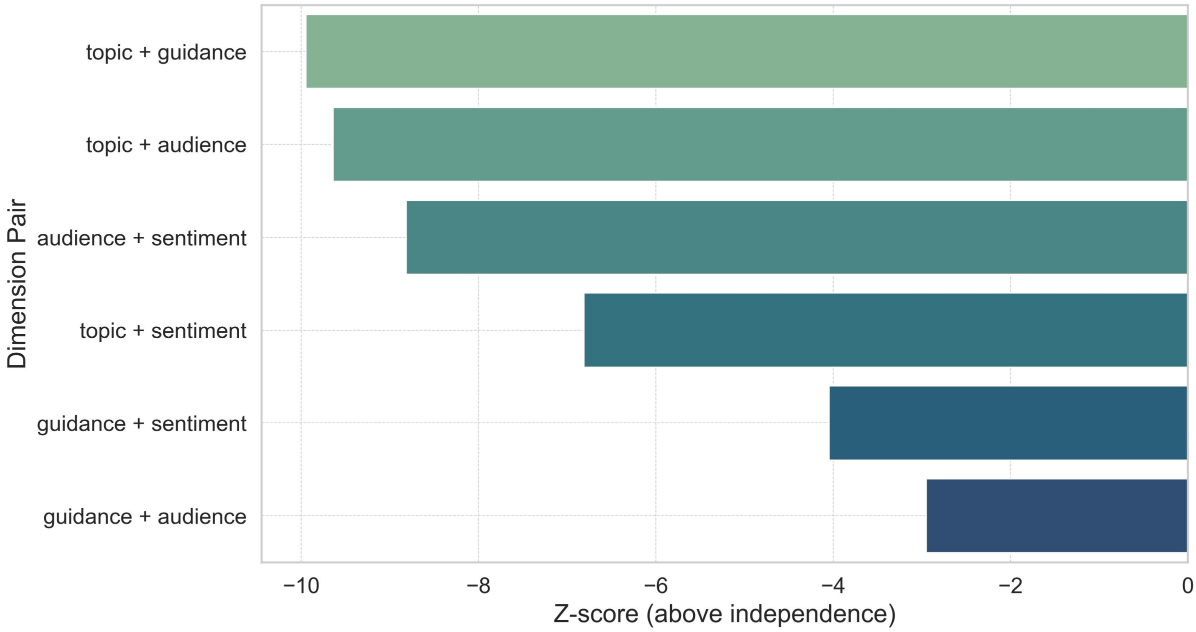
Figure 7 presents the taxonomy of topics and communication stances used in the first classifier. We design a list of topics to comprehensively cover the potential range of central bank discourse, encompassing monetary policy, financial stability, supervision and regulation, payments, and structural economic issues. Additional granularity is introduced in the monetary policy domain by further subdividing it into key subtopics, such as interest rates, inflation, and balance sheet size, including asset purchase programs. This finer categorization is particularly important for the empirical analysis, as it allows us to examine how central banks adjust their messaging when transitioning between different monetary policy frameworks (e.g., adopting inflation targeting).

To ensure consistency when labeling the sentence topic, we classify statements based on their economic origin rather than their effect. This approach provides greater objectivity, as the underlying driver of a statement tends to be uniquely identifiable, whereas its effects may span multiple domains. For example, in the sentence "Monetary policy tightening can significantly raise borrowing costs, potentially putting pressure on the stability of financially constrained firms," the primary origin is monetary policy (interest rate decisions), even though financial stability considerations are mentioned. Thus, the statement is classified under Monetary Policy (Interest Rate) rather than Financial Stability.

We categorize the communication stance as backward-looking or forward-looking. Backward-looking statements provide assessments of past or current economic conditions or policy decisions, while forward-looking statements offer projections, guidance, or expectations about future policy actions and financial developments. Significantly, the stance is not determined

[7] To quantify co-occurrence structure beyond chance, we compute a Z-score for each pair of label dimensions. Specifically, for each pair (e.g., *topic* and *guidance*), we calculate the number of unique label combinations observed in the dataset. We then construct a null distribution by fixing one label and randomly permuting the other 10,000 times, recording the number of unique combinations in each permutation. The Z-score is computed as $Z = (C_{\text{obs}} - \mathbb{E}[C_{\text{rand}}])/\text{SD}[C_{\text{rand}}]$, where C_{obs} is the observed number of unique combinations, $\mathbb{E}[C_{\text{rand}}]$ and $\text{SD}[C_{\text{rand}}]$ are the expected number and associated standard deviation under permutation. While the highest absolute Z-scores were associated with the pairs *topic* + *guidance* and *topic* + *audience*, only the former could be selected due to overlap in the topic dimension. Therefore, we chose the pair *audience* + *sentiment* as the second classifier to maximize the total co-occurrence signal while ensuring disjoint dimensional coverage.

Figure 6: Empirical Pairwise Coupling Between Classification Dimensions



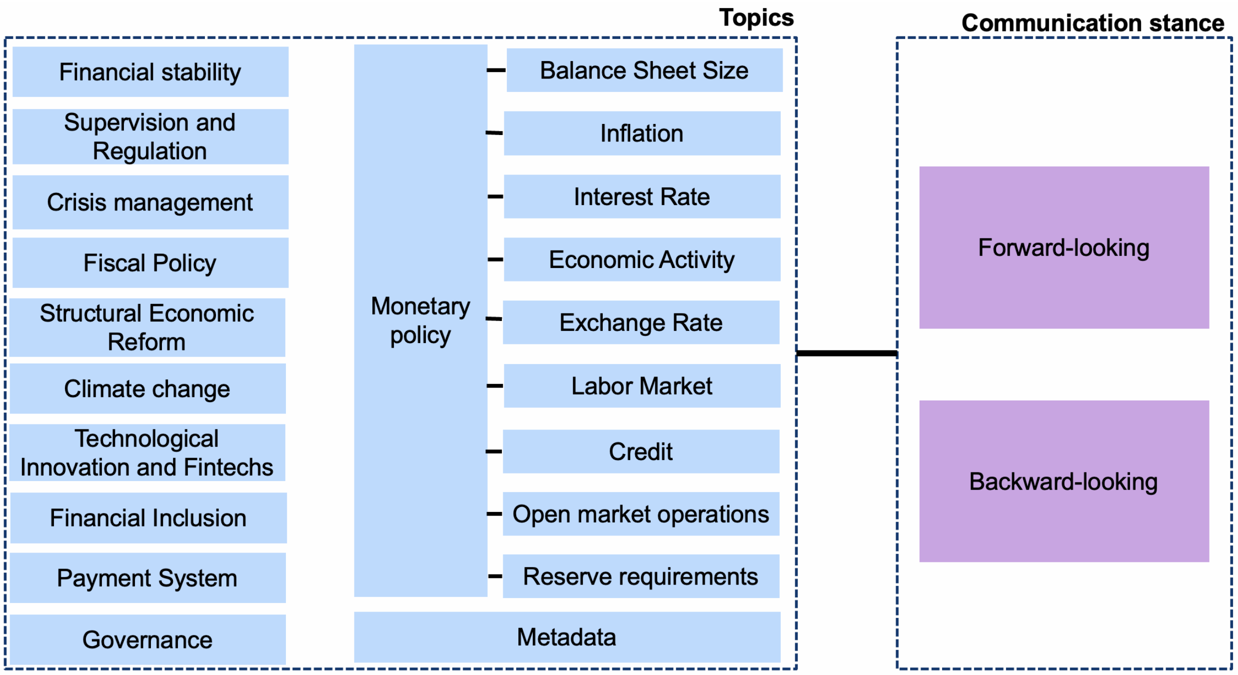
Notes: This figure reports the empirical pairwise coupling between the four classification dimensions using Z-scores. The figure ranks all pairs of sentence-level labels—topic, communication stance, audience, and sentiment, based on how strongly they co-occur relative to what would be expected under statistical independence. Each Z-score compares the number of unique co-occurrences observed in the labeled dataset against a null distribution generated by randomly permuting one label while holding the other fixed. Higher absolute Z-scores indicate greater empirical structure. The pairs *topic* + *guidance* and *audience* + *sentiment* show the strongest dependency, supporting their use as the basis for the two jointly estimated classifier modules. The null ($Z = 0$) represents the threshold for statistical independence.

solely by verb tense; rather, it reflects whether the message's substantive content pertains to future developments. For instance, the sentence "Last year's findings emphasized the need for continued efforts to address disparities in access to banking services" is classified as Financial Inclusion (Forward-Looking), despite referencing past findings, because the primary message concerns prospective action. This distinction is essential for accurately capturing the policy signaling embedded in central bank communications.

Jointly estimating topic and communication stance is motivated not only by their empirical co-occurrence structure but also by the economic logic of central bank discourse. The substantive content of a message (e.g., inflation, financial stability) and its temporal orientation (e.g., projection, assessment) are conceptually interdependent: forward guidance on interest rates, for instance, differs fundamentally in tone and implications from *ex-post* justification of the same policy. Therefore, explicitly modeling this interaction has predictive power that our empirical setup can explore.

The second classifier jointly predicts audience and sentiment. This setup is flexible to capture how central banks tailor their communication to different stakeholder groups while modulating tone accordingly. Figure 8 displays our classification framework's taxonomy of audiences and

Figure 7: Classes for the Topic and Communication Stance Dimensions

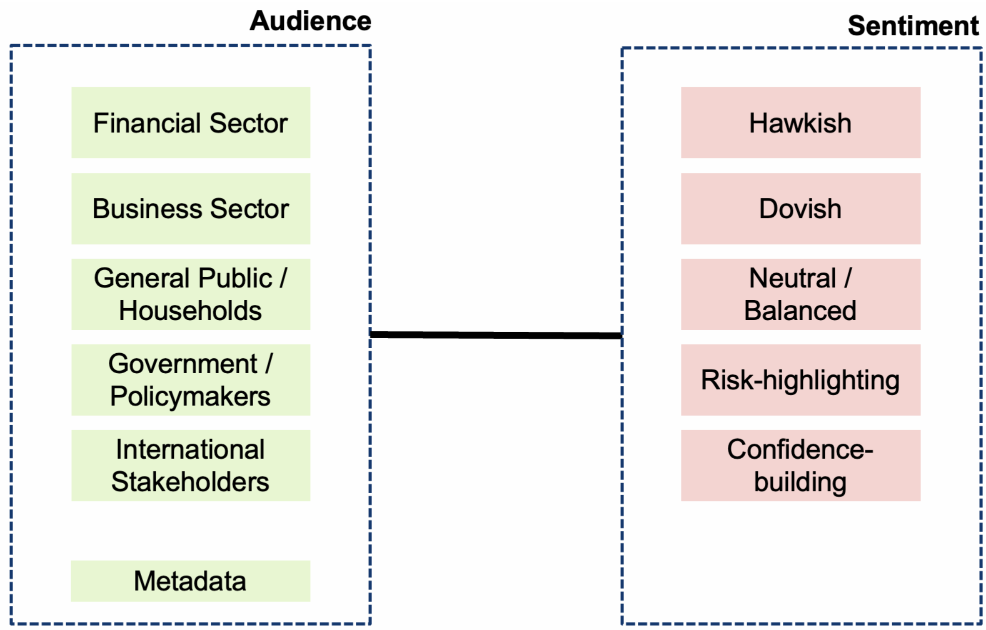


Notes: This figure schematizes the set of classes considered for the topic and communication stance dimensions in the classification framework at the sentence level. The classification nests both dimensions, meaning there is differentiation between forward- and backward-looking communication stances for the same topic.

sentiments. Economically speaking, jointly estimating these two dimensions reflects the strategic nature of policy communication: central banks adapt their tone, complexity, and rhetorical stance depending on whether they address financial markets, businesses, government officials, households, or international partners. For example, statements directed at financial markets are typically precise and data-driven, signaling policy intent through subtle shifts in language. In contrast, statements aimed at the general public are more likely to be reassuring or simplified, as they help anchor expectations and maintain trust. Therefore, modeling sentiment and audience jointly captures a key layer of policy signaling that varies systematically with the intended recipient.

For both classifiers, we introduce an additional class labeled "Metadata," which captures sentences that do not convey economic meaning. This class includes data source references in figure captions, formalities in opening and closing remarks, acknowledgments, boilerplate disclaimers, and procedural statements such as "The Monetary Policy Committee approved this report on [date]." We include representative examples of this residual class in our labeled dataset to ensure proper classification. This approach offers a more robust and elegant filtering mechanism compared to methods that discard sentences based on keywords or length, as it removes only those semantically unrelated to economic content, thereby reducing the risk of mistakenly excluding relevant statements. In the final classified results, the "Metadata" category comprised approximately 6.3 percent of the total for topic and communication stance classification, and 14.1 percent for audience and sentiment classification.

Figure 8: Classes for the Audience and Sentiment Dimensions



Notes: This figure schematizes the set of classes considered for the audience and sentiment dimensions in the classification framework at the sentence level. The classification nests both dimensions, allowing for differentiation between the same sentiment across audiences.

4.1.4. Fine-Tuning Setup

The fine-tuning process follows two sequential phases (Tunstall et al., 2022). First, a Siamese neural network learns fixed-length, dense vector representations of sentences. In the case of the chosen bge-m3 model, this is a 1024-dimensional vector space. In this setup, the model processes sentence pairs rather than individual sentences, optimizing a contrastive learning loss function that brings semantically similar pairs closer in the embedding space while pushing dissimilar ones apart. Second, we train a dense neural network to map these learned representations to classification labels. A key advantage of contrastive learning is that it effectively expands the training set size, as working with sentence pairs instead of isolated examples yields up to $\frac{N(N-1)}{2}$ training pairs from a dataset of size N.[8] This is particularly valuable given the challenges of assembling high-quality labeled datasets, which require significant expert input and manual annotation.

Importantly, the generation of up to $N(N-1)/2$ pairs does not aim to introduce new informational content *per se*, but to guide the model in shaping a semantically meaningful embedding space. By explicitly modeling similarities and differences between sentence pairs, contrastive learning provides richer optimization signals that help the model learn more discriminative features than possible with isolated instances. In other words, the benefit comes not from additional observations, but from the richer optimization signals that emerge when

[8] Even for small N, this approach generates a sufficiently large dataset for effective learning. Our training dataset, comprising 840 examples, would correspond to 352,380 pairs if no up-sampling were done.

relationships between sentence pairs are explicitly modeled. This is a well-established advantage of contrastive and metric learning methods, particularly in low-resource settings or when class boundaries are semantically subtle (Khosla et al., 2020).

The fine-tuning process follows two consecutive steps. In the first step, we fine-tune the sentence transformer to produce embedding vectors using the cosine similarity metric for distance calculation, which is well-suited for textual data. The contrastive loss function is a hard triplet loss, which numerically encourages the model to distinguish between similar and dissimilar sentences. During fine-tuning, we create all possible sentence pair combinations within each labeled dataset without oversampling or undersampling, as the class distribution is already balanced.

We perform hyperparameter tuning on two critical parameters: the learning rate, searched within the range $[10^{-5}, 10^{-2}]$, and the L_2-norm regularization, searched within $[10^{-6}, 10^{0}]$. We employ a variant of the Adam optimizer that decouples weight decay from the adaptive gradient updates, thereby enhancing generalization in deep learning models. In our training, the numerical optimizer uses a learning rate scheduler incorporating a warm-up phase (10 percent of an epoch) followed by decay, allowing for stable initial convergence and gradual refinement as training progresses. Every 500 steps, we evaluate the model's out-of-sample performance on the validation set, optimizing for the embedding loss, which provides a meaningful measure of how well the sentences are positioned within the semantic space. We keep track of the model that minimizes the loss function during the training phase.

In the second step, we apply a softmax function in the output layer (after the dense layer), with the number of neurons corresponding to the number of classes. We perform a model selection procedure similar to the first part. However, we optimize for out-of-sample accuracy on the validation set, which ensures that the model generalizes well to unseen data.

4.1.5. Out-of-Sample Performance

We benchmark the performance of our classification framework against ChatGPT 4o, a state-of-the-art general-purpose language model. This exercise utilizes our small validation set, where costs are manageable.[9] While not explicitly trained for classification tasks, generative large language models like ChatGPT have demonstrated strong zero- and few-shot capabilities across a

[9] We estimate the cost of processing our entire dataset of approximately 21 million central bank communication sentences using OpenAI's models, based on pricing as of April 2025. GPT-4o is priced at \$2.50 per million input tokens and \$10.00 per million output tokens, while GPT-4.5 is significantly more expensive at \$75.00 per million input tokens and \$150.00 per million output tokens. Assuming each prompt contains 250 words (approximately 333 tokens) and each response contains 10 words (approximately 13 tokens), a single prompt–response interaction would cost \$0.00096 with GPT-4o and \$0.02693 with GPT-4.5 (USD). Processing 21 million such interactions would therefore cost approximately \$20,160 with GPT-4o and \$565,530 with GPT-4.5. One might consider batching multiple sentences into a single prompt to reduce costs. While this approach can decrease the number of API calls, it is still not scalable. Also, multiple sentences in the same prompt increase the propensity for hallucinations, potentially leading to inaccurate or fabricated outputs. Moreover, any change in the classification schema—such as adding a new topic or modifying label definitions—would require reprocessing the entire dataset to maintain consistency, compounding both computational and financial costs. Finally, outputs generated by commercial LLMs such as ChatGPT are inherently non-reproducible and non-transferable across users or time, limiting their use in research and policy applications, as pointed out for example in Gambacorta et al. (2024). By contrast, our trained classification model is reproducible, version-controlled, and can be shared with others to ensure transparent and consistent results.

wide range of natural language understanding problems. Evaluating ChatGPT on our classification task provides a meaningful reference point for assessing the value of domain-specific fine-tuning relative to a highly capable, commercially available alternative.

We adopt a weakly supervised evaluation protocol in which ChatGPT 4o is instructed to assign one label for each of the four sentence-level dimensions—topic, communication stance, audience, and sentiment—based on a fixed set of allowable classes. See Prompt A.2 for the detailed prompt. If an invalid response is returned (i.e., outside the specified classes), the query is repeated until a valid output is obtained. This evaluation simulates a realistic usage scenario in which an analyst leverages ChatGPT for structured annotation within predefined categories.

Table 2 compares our classifier's performance on the validation set against ChatGPT 4o across topic, communication stance, audience, and sentiment. Our classifier consistently achieves higher macro-level metrics, notably macro F1 scores, reflecting superior performance across minority classes. The most substantial performance gaps occur in the communication stance and audience dimensions, where our model significantly outperforms ChatGPT 4o. Central banks carefully frame their messages with precise temporal stances (forward- versus backward-looking) and tailor their tone and complexity according to specific audiences. These nuanced, economically informed distinctions require targeted fine-tuning, which our domain-specific classifier explicitly incorporates. In contrast, ChatGPT 4o, as a general-purpose generative model, lacks the tailored optimization needed to systematically capture such subtle semantic differences, causing it to default disproportionately to frequent or broadly generalizable classes.

Table 2: Comparative Performance of Our Classifier versus ChatGPT 4o

Dimension	**Model**	**Accuracy**	**Precision (Macro)**	**Recall (Macro)**	**F1 (Macro)**	**Precision (Micro)**	**Recall (Micro)**	**F1 (Micro)**	**Cohen's Kappa**
Topic	Our Classifier	0.689	0.699	0.645	0.650	0.689	0.689	0.689	0.666
	ChatGPT 4o	0.731	0.593	0.535	0.551	0.731	0.731	0.731	0.711
Comm. stance	Our Classifier	0.924	0.925	0.910	0.917	0.924	0.924	0.924	0.834
	ChatGPT 4o	0.828	0.591	0.551	0.570	0.828	0.828	0.828	0.658
Audience	Our Classifier	0.706	0.713	0.699	0.700	0.706	0.706	0.706	0.622
	ChatGPT 4o	0.506	0.529	0.463	0.437	0.506	0.506	0.506	0.400
Sentiment	Our Classifier	0.700	0.612	0.595	0.594	0.700	0.700	0.700	0.589
	ChatGPT 4o	0.704	0.557	0.393	0.418	0.704	0.704	0.704	0.592

Notes: Accuracy is the proportion of correctly classified instances across all classes. *Precision* is the proportion of true positives among predicted positives, measuring reliability of positive predictions. *Recall* is the proportion of true positives among actual positives, measuring completeness. *F1-score* is the harmonic mean of precision and recall, providing a balanced measure of accuracy. Metrics marked *Macro* compute the arithmetic mean across all classes, giving equal importance to each class regardless of size. Metrics marked *Micro* aggregate true positives, false positives, and false negatives across all classes, thus being weighted by class frequency and reflecting performance on more frequent categories. *Cohen's Kappa* corrects accuracy for chance agreement, providing robustness to class imbalance.

Regarding the sentiment dimension, while both models achieve similar accuracy and micro-F1 scores, our classifier attains considerably higher macro-F1 scores. This indicates better performance on less prevalent sentiment categories, such as hawkish or dovish tones, which have significant implications for market expectations. The observed discrepancy underscores a known limitation of general-purpose generative models: their outputs tend toward frequent or neutral classes, often obscuring subtle yet critical signals of economic policy sentiment. Our specialized approach, fine-tuned specifically for central bank discourse, mitigates this limitation by capturing nuanced shifts across the sentiment spectrum more effectively.[10]

For the topic dimension, ChatGPT 4o achieves slightly higher accuracy and micro-F1 due to strong performance on frequent classes. However, it suffers a markedly lower macro-F1 score compared to our model. This divergence reflects ChatGPT's difficulty handling infrequent yet economically meaningful topics, which are critical in monitoring policy framework shifts or new central bank attention areas (e.g., digital currencies, climate risks). By contrast, our model, trained specifically with domain expertise and comprehensive economic classification frameworks, provides more balanced and robust classification across both mainstream and less frequent policy topics.

We also performed an error analysis of the output to check for structural inconsistencies. Many of the misclassifications are justifiable upon closer inspection, as the distinctions between related categories can be subtle. For instance, a few sentences on "monetary policy – open market operations (forward-looking)" were misclassified as "monetary policy – inflation (forward-looking)" and "monetary policy – balance sheet size (forward-looking)." In these instances, open market operations citations were used as monetary policy tools to influence short-term interest rates and liquidity in the financial system, aiming to achieve price stability. Similarly, some few sentences on "fiscal policy (backward-forward)" were misclassified as "monetary policy – economic activity (backward-looking)" and "monetary policy – exchange rate (backward-looking)." In these, fiscal policy was brought about to influence economic activity and exchange rates. Lastly, some sentences on "financial stability (backward-looking)" were misclassified as "supervision and regulation (backward-looking)" due to the overlap between regulatory actions and financial stability outcomes. These errors are understandable given the inherent overlap in the semantic content of the classes, where fiscal, monetary, and regulatory policies often intersect and influence one another.

4.2. Empirical Application on the Central Bank Communications Dataset

This section applies the fine-tuned classifier to the large dataset of central bank communications. Our unit of analysis is the sentence in a particular document. To get a sense of the classifier's output, Table 3 distills the ECB's monetary policy decision published on December 12, 2024, and shows the classifier's output for the four dimensions discussed above. In general, each document type we collected in Table 1 has a different publication frequency. For example, monetary policy decisions are released several times a year, while annual reports are published once a year. This

[10] Bucher & Martini (2024) show that fine-tuned, task-specific models consistently outperform larger, zero-shot generative models like ChatGPT in text classification tasks, particularly in handling less frequent classes. The authors highlight that generative models often default to more common or neutral classes, which can mask poor performance on minority or nuanced categories. This finding supports the assertion that a specialized approach, fine-tuned for central bank discourse, captures nuanced shifts across the sentiment spectrum more effectively.

section aggregates data at the semiannual level to facilitate visual inspections, using within-year interpolation when necessary for less frequent documents, such as annual reports.[11]

Table 3: Classification of the ECB's Monetary Policy Decision Published on December 12, 2024

Sentence	Topic	Comm. Stance	Audience	Sentiment
PRESS RELEASE Monetary policy decisions 12 December 2024 The Governing Council today decided to lower the three key ECB interest rates by 25 basis points.	MP - interest rate	Backward-looking	Financial Sector	Dovish
In particular, the decision to lower the deposit facility rate – the rate through which the Governing Council steers the monetary policy stance – is based on its updated assessment of the inflation outlook, the dynamics of underlying inflation and the strength of monetary policy transmission.	MP - interest rate	Forward-looking	Financial Sector	Dovish
The disinflation process is well on track.	MP - inflation	Forward-looking	Financial Sector	Confidence-building
Most measures of underlying inflation suggest that inflation will settle at around the Governing Council's 2% medium-term target on a sustained basis.	MP - inflation	Forward-looking	Financial Sector	Neutral / Balanced
Domestic inflation has edged down but remains high, mostly because wages and prices in certain sectors are still adjusting to the past inflation surge with a substantial delay.	MP - inflation	Backward-looking	General Public	Risk-highlighting
Financing conditions are easing, as the Governing Council's recent interest rate cuts gradually make new borrowing less expensive for firms and households.	MP - interest rate	Forward-looking	Business Sector	Dovish
But they continue to be tight because monetary policy remains restrictive and past interest rate hikes are still transmitting to the outstanding stock of credit.	MP - interest rate	Backward-looking	Business Sector	Risk-highlighting
Staff now expect a slower economic recovery than in the September projections.	MP - economic activity	Forward-looking	Business Sector	Neutral / Balanced
Although growth picked up in the third quarter of this year, survey indicators suggest it has slowed in the current quarter.	MP - economic activity	Backward-looking	Business Sector	Risk-highlighting

[11] For example, if a central bank publishes an annual report in June 2020, the same value is carried forward and used for both the first and second halves of 2020 in the aggregation. This approach ensures consistency in time series coverage across document types with different publication frequencies.

Table 3 (continued)

Sentence	Topic	Comm. Stance	Audience	Sentiment
The projected recovery rests mainly on rising real incomes—which should allow households to consume more—and firms increasing investment.	MP - economic activity	Forward-looking	General Public	Confidence-building
Over time, the gradually fading effects of restrictive monetary policy should support a pick-up in domestic demand.	MP - interest rate	Forward-looking	General Public	Dovish
The Governing Council is determined to ensure that inflation stabilises sustainably at its 2% medium-term target.	MP - inflation	Forward-looking	Financial Sector	Neutral / Balanced
It will follow a data-dependent and meeting-by-meeting approach to determining the appropriate monetary policy stance.	MP - interest rate	Forward-looking	Financial Sector	Confidence-building
The Governing Council is not pre-committing to a particular rate path.	MP - interest rate	Forward-looking	Financial Sector	Neutral / Balanced
The Eurosystem no longer reinvests all of the principal payments from maturing securities purchased under the PEPP, reducing the PEPP portfolio by €7.5 billion per month on average.	MP - balance sheet	Backward-looking	Financial Sector	Hawkish
The Governing Council will discontinue reinvestments under the PEPP at the end of 2024.	MP - balance sheet	Forward-looking	Financial Sector	Hawkish
The Governing Council stands ready to adjust all of its instruments within its mandate to ensure that inflation stabilises sustainably at its 2% target over the medium term and to preserve the smooth functioning of monetary policy transmission.	MP - inflation	Forward-looking	Financial Sector	Confidence-building
Moreover, the Transmission Protection Instrument is available to counter unwarranted, disorderly market dynamics that pose a serious threat to the transmission of monetary policy across all euro area countries, thus allowing the Governing Council to more effectively deliver on its price stability mandate.	Governance	Forward-looking	Financial Sector	Confidence-building
The President of the ECB will comment on the considerations underlying these decisions at a press conference starting at 14:45 CET today.	Metadata	Metadata	Metadata	Metadata
CONTACT European Central Bank Directorate General Communications Sonnemannstrasse 20 60314 Frankfurt am Main, Germany +49 69 1344 7455 media@ecb.europa.eu Reproduction is permitted provided that the source is acknowledged.	Metadata	Metadata	Metadata	Metadata

In interpreting the aggregated outputs of the classifier across central banks, it is essential to establish the reliability of sentence-level predictions. Appendix B shows that the classification framework performs consistently across languages, with only minor differences observed between original and translated documents. This ensures that multilingual publications can be analyzed jointly without introducing material distortions. Furthermore, Appendix C confirms that predictions are typically made with high confidence. The classifier assigns unambiguous labels in the large majority of cases. Finally, Appendix D indicates that sentences involving multiple overlapping classifications are relatively rare. For example, instances where two topics or sentiments are simultaneously predicted occur infrequently and do not materially affect the interpretation of results. Taken together, these findings reinforce that the sentence-level classifications used in the subsequent analysis are linguistically robust, highly confident, and predominantly unambiguous, supporting their aggregation across countries and over time.

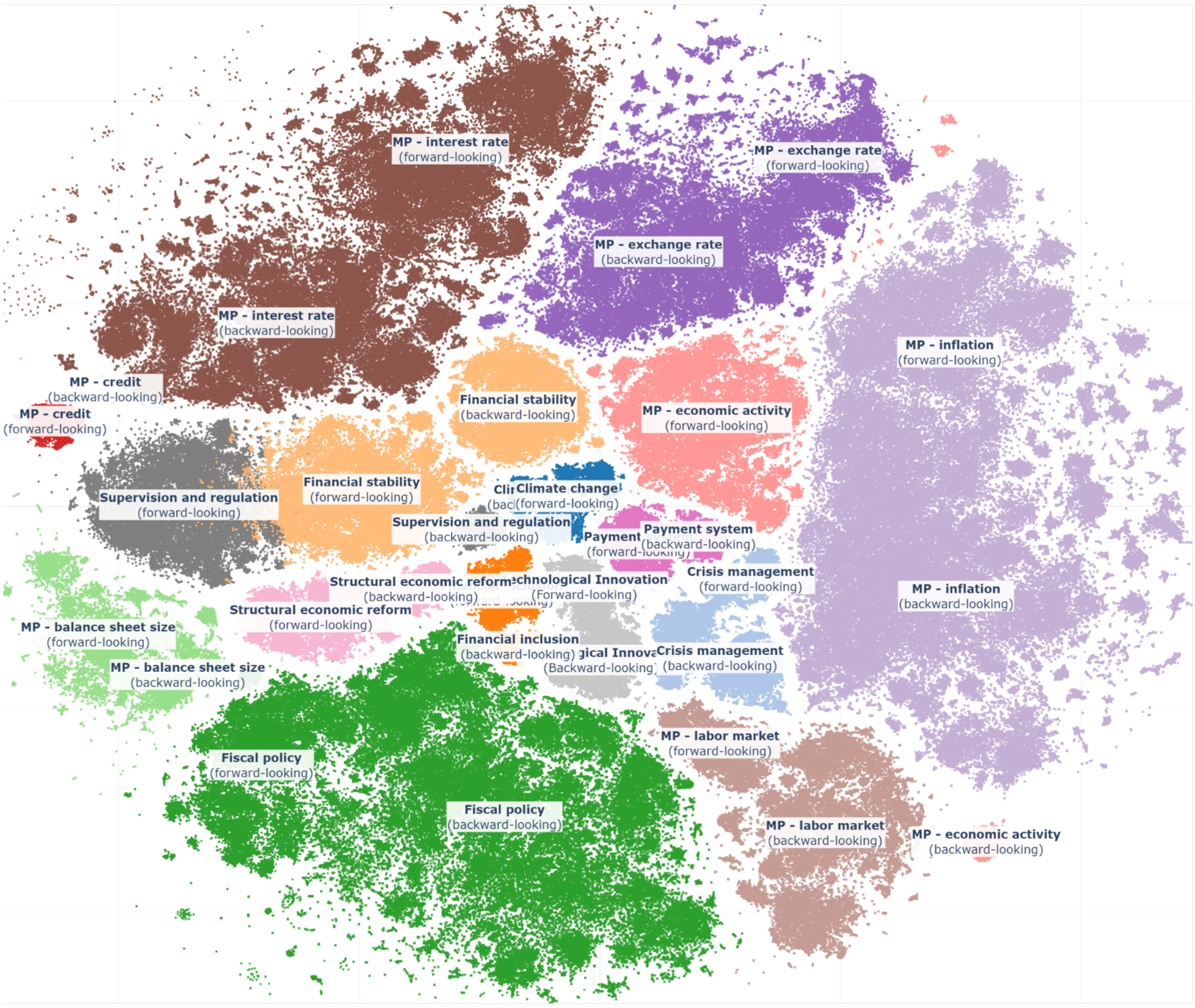
Figure 9 presents a semantic map of central bank communications in the database from 1884 to 2025, offering a structured visual representation of how different topics and communication stances relate over a century of central bank publications. We reduce the original 1024-dimensional embedding space to a two-dimensional representation using the t-SNE nonlinear dimensionality reduction technique (van der Maaten & Hinton, 2008), which preserves local and global semantic structures more effectively than linear alternatives such as PCA (Silva & Zhao, 2016). The reduction is performed in an unsupervised way, i.e., no class labels were used except for visualization purposes at the end. Each dot represents a sentence, with proximity indicating semantic similarity. The visualization captures the thematic structure of central bank discourse.

There are clear clustering patterns, suggesting well-defined topic boundaries. A notable exception is the overlap between forward-looking supervision and regulation and financial stability, an expected outcome given their shared focus on systemic risk and regulatory oversight. Forward- and backward-looking statements within the same topic typically appear adjacent, reflecting their temporal alignment while maintaining semantic coherence.

The cluster width represents semantic variation. The map is structured to preserve both intra- and interclass similarity. "Traditional" topics, such as inflation, fiscal policy, exchange rates, and interest rates, exhibit broad semantic variation, highlighting their complexity, the diverse contexts in which they are discussed, and their distinct linguistic and conceptual characteristics. In contrast, "emerging" topics, such as climate change and technological innovation, show less dispersion on the semantic map, indicating that similar types of sentences tend to repeat more for "emerging" than for "traditional" topics. Furthermore, classes positioned in the middle, such as emerging topics, tend to be semantically more similar to all other classes. This arrangement could be explained by central banks' efforts to reference "traditional" topics when addressing "emerging" ones. This semantic mapping offers valuable insights into the evolution of central bank communication, highlighting how certain topics function as specialized domains. In contrast, others serve as integrative themes within the broader economic discourse.

Figure 10 exhibits the global evolution of central bank communication topics across different communication outlets, highlighting variations in thematic focus over time. As expected, monetary policy is the dominant topic in central banks' decisions and reports, aligning with their core price

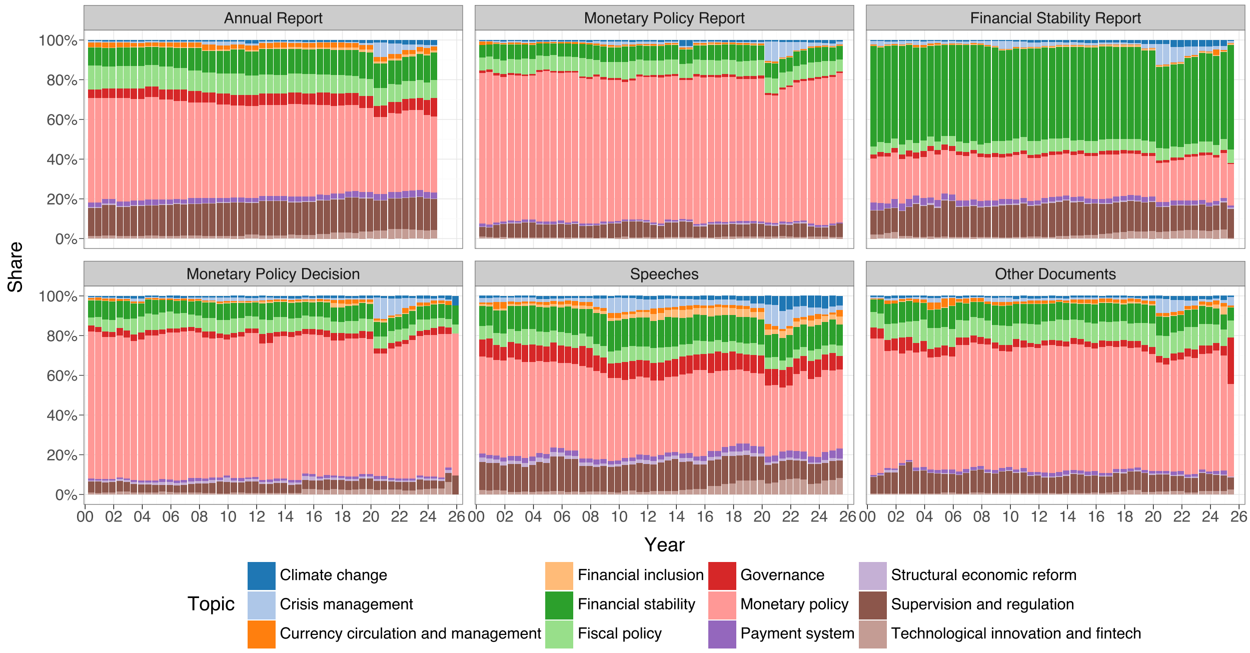
Figure 9: Visual Representation of Central Bank Communication Over a Century of Data



Notes: This figure displays the semantic space of central bank communications classified by topic and communication stance. Each dot represents a sentence, with its position reflecting semantic similarity to others. Colors indicate different topics, each containing both backward- and forward-looking sentences. The original 1024-dimensional sentence embeddings are projected onto two dimensions using the t-SNE nonlinear dimensionality reduction technique (van der Maaten & Hinton, 2008) in an unsupervised way, i.e., no class labels were used except for visualization purposes at the end.

stability mandate. Similarly, financial stability reports allocate the most attention to financial stability, supervision, and regulation, reflecting their institutional role in monitoring systemic risk. While crisis management and fiscal policy exhibit episodic spikes, often in response to economic and financial crises, emerging topics such as climate change and technological innovation have gained space in most documents in recent years. Speeches display the most remarkable thematic diversity, dedicating more attention to emerging issues than other communication outlets. Speeches provide a flexible platform for addressing evolving challenges beyond traditional monetary and

Figure 10: Topic Composition of Central Bank Communications by Communication Outlet



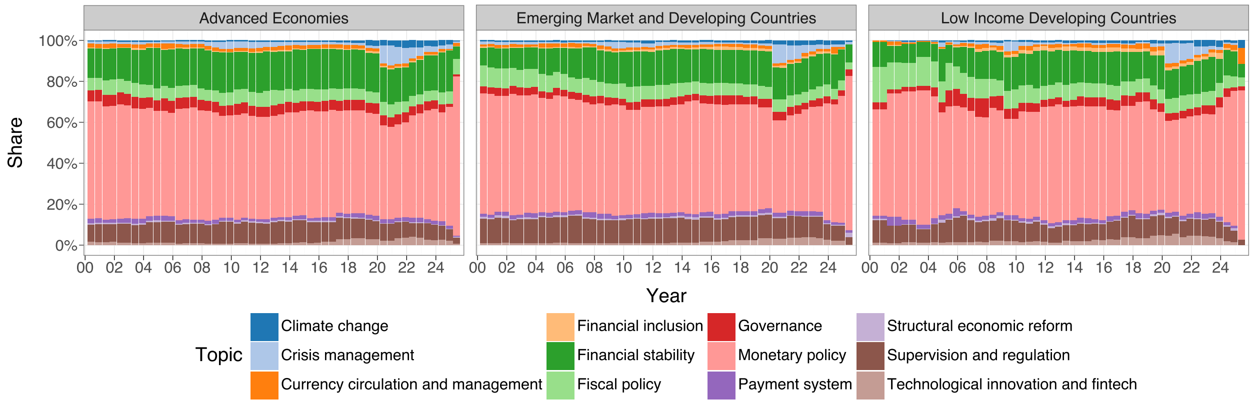
Notes: This figure shows the global evolution of central bank communication topics across different communication outlets. Colors indicate the topic. For each document, we evaluate the number of sentences in a specific topic as a share of the total number of sentences. The figure shows the average share evaluated across documents of the same type published by all central banks in the same semiannual period. The monetary policy topic encompasses all the subtopics discussed in Figure 7.

financial stability concerns.

Figure 11 breaks down the topic distributions by the level of economic development. The level of development classification for each economy is taken from the IMF AREAER dataset (International Monetary Fund, 2025). Interestingly, the suite of topics central banks discuss is remarkably similar across economies of different market types. However, notable differences emerge in the relative emphasis placed on specific topics. Advanced economies devote more attention to financial stability, reflecting their relatively more sophisticated financial markets and the need for prudential oversight. In contrast, emerging and low-income economies place greater importance on fiscal policy, particularly in low-income countries where fiscal-monetary interactions are more pronounced due to limited market depth and a reliance on central bank financing of the government. These differences highlight how structural and institutional factors shape the communication priorities of central banks across economies at various stages of development.

We observe significant heterogeneity when focusing on central bank communication at the economy level. Figure 12 illustrates the topic distribution for the United Kingdom and the United States, two economies with the most extensive available time series of central bank communications. While central bank communication patterns have generally remained consistent over the last 25 years, significant shifts occurred in earlier periods. One notable exception is the monetary policy topic, which has exhibited a relatively stable distribution across nearly a century of data, reflecting

Figure 11: Topic Composition of Central Bank Communications by Level of Development



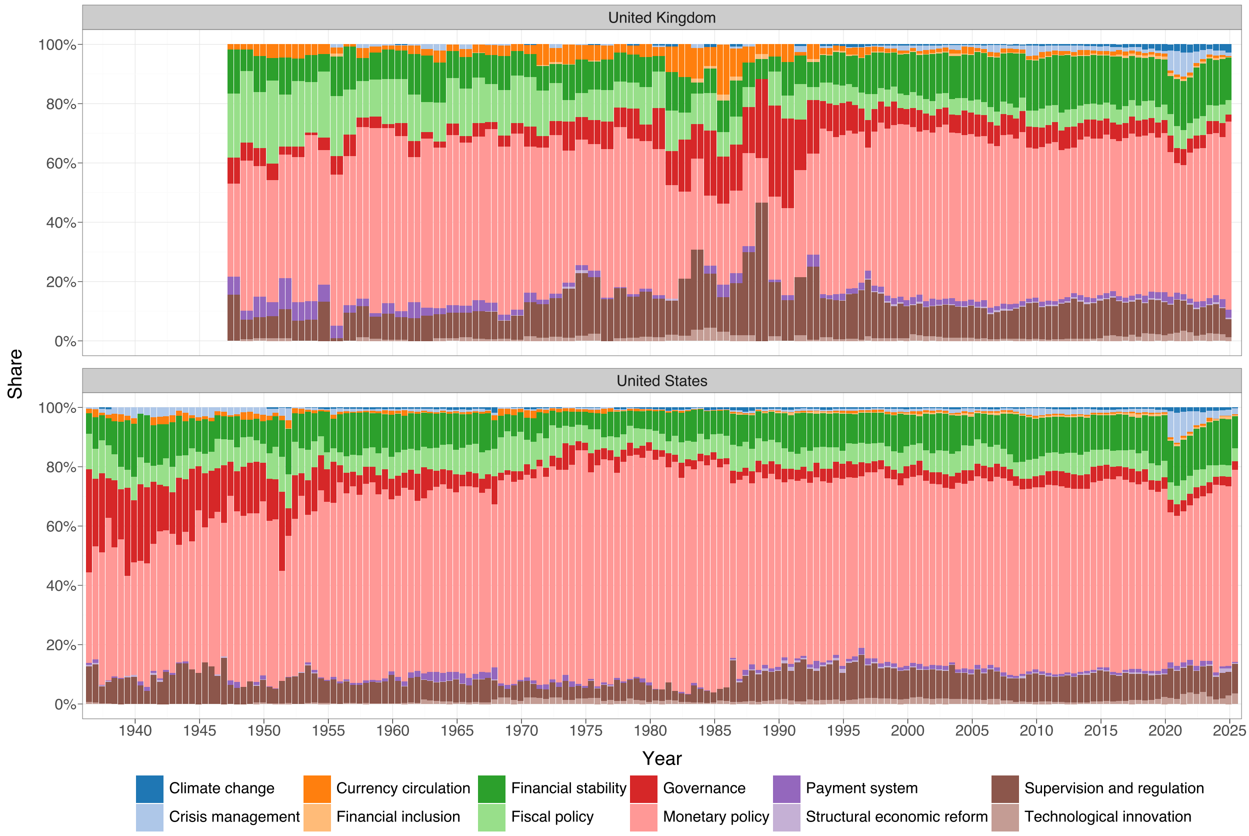
Notes: This figure shows the evolution of central bank communication topics by level of economic development: advanced economies (left), emerging market and developing countries (center), and low-income developing countries (right). Colors indicate the topic. For each document, we evaluate the number of sentences in a specific topic as a share of the total number of sentences. The figure shows the average share evaluated across documents of central banks within economies of the same market type in the same semiannual period. The monetary policy topic encompasses all the subtopics discussed in Figure 7.

its central role in the mandates of central banks. However, specific topics surged at key historical junctures. The UK's governance topic gained notable attention between 1985 and 1990, reflecting the government's efforts to enhance central bank independence. In the United States, governance statements were prominent from 1936 to 1955, coinciding with the development of the Federal Reserve's modern framework after the Great Depression and its critical role in managing wartime economic policy. Crisis management gained prominence in the US between 1940 and 1946, driven by the economic turmoil of World War II, which necessitated coordinated fiscal and monetary measures. The rise of financial stability concerns in the 1990s in both countries coincided with the implementation of the Basel Accord, the growing recognition of systemic risk, and the increasing interconnectedness of the global financial system. Note that all of these features are results of the classification, which did not explicitly mention any of these topics, and this verifies that our classification produces sensible results on historical and cross-country data spanning decades.

We now consider topics on monetary policy communication only. We break it down into the subtopics and communication stances listed in Figure 7. Figure 13 displays the shares of each monetary policy subtopic and communication stance by the level of economic development.[12] Advanced economies focus on signaling their monetary policy stance, as reflected in the prominent share of communication dedicated to interest rates. Their communication is also more forward-looking, focusing on signaling future monetary policy actions. In contrast, emerging and low-income economies prioritize inflation over interest rates, likely due to ongoing efforts to anchor inflation expectations while introducing regimes that target inflation more directly. A notable trend across all economies is the declining emphasis on exchange rates. This fact can be attributed to the increased adoption of inflation-targeting regimes, which have reduced reliance on exchange

[12] That is, we are only considering the pink bars displayed in Figure 10.

Figure 12: Topic Composition of Central Bank Communications in the United Kingdom and United States

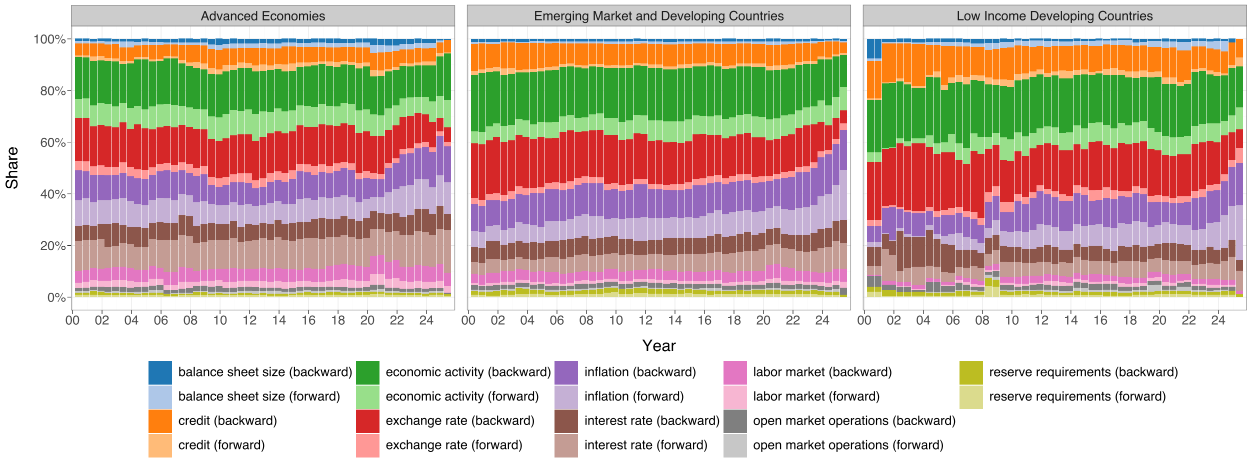


Notes: This figure shows the evolution of central bank communication topics for the United Kingdom and the United States. For each document, the share of sentences classified under each topic is computed, and semiannual averages are shown. Colors represent different topics. The monetary policy topic encompasses all the subtopics discussed in Figure 7.

rate interventions as a monetary policy tool. However, discussions on exchange rates remain more relevant in low-income economies, where currency stability and external vulnerabilities are central concerns due to their exchange rate arrangements.

Figure 14 shows the evolution of monetary policy communication in the United Kingdom and the United States. Before adopting inflation targeting (IT) in the UK, communication was mainly backward-looking, focusing on exchange and interest rates. Following the adoption of IT, the UK experienced a shift towards more forward-looking statements, particularly regarding interest rates, inflation, and economic activity, reflecting the forward-looking nature of the IT framework. Similarly, in the USA, forward-looking communication has increased over time, with a continued emphasis on economic activity and the labor market. While both countries have become more forward-looking in their monetary policy communication, a key difference is that the US maintains a stronger focus on economic activity and labor market conditions when conducting monetary policy communication. In contrast, the UK devotes more statements to inflation and interest rates. Additionally, both countries have reduced their discussion on monetary policy tools, such as

Figure 13: Forward- and Backward-Looking Monetary Policy Communication by Level of Development



Notes: This figure shows the evolution of monetary policy communication across economies with different levels of development: advanced economies (left), emerging market and developing countries (center), and low-income developing countries (right). Colors indicate the monetary policy subtopic. Darker colors represent backward-looking shares, while lighter colors represent forward-looking shares. For each document, we evaluate the number of sentences in a specific subtopic as a share of the document's total number of monetary policy sentences. The figure presents the average share across all documents of the same type published by economies with a similar level of development.

open market operations and reserve requirements, in their official monetary policy communication outlets.

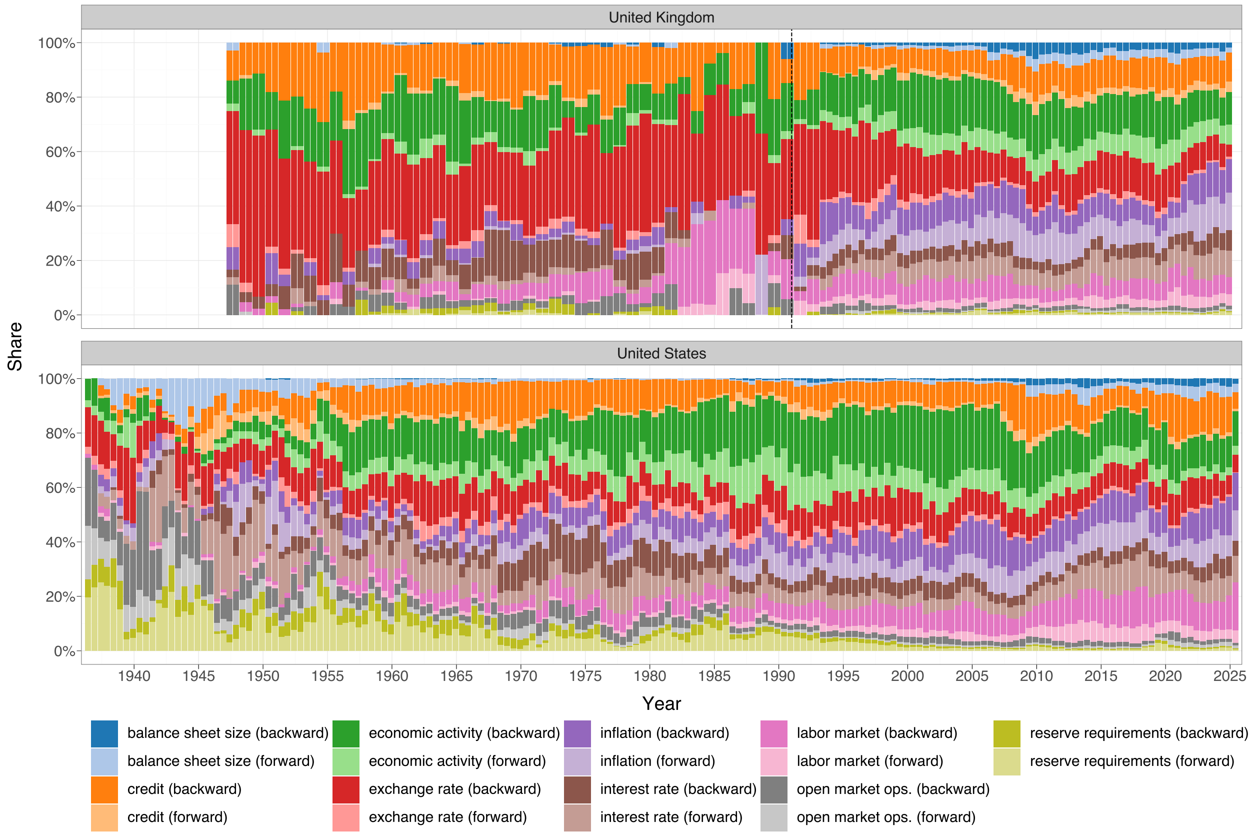
IT adoption leads to structural shifts in monetary policy communication, and a common pattern emerges across economies, regardless of their level of development. Figure 15 shows the same information as above but for Brazil, Chile, Georgia, Republic of Kazakhstan, Republic of Korea, Mexico, Republic of Moldova, New Zealand, Paraguay, Peru, the Philippines, Russian Federation, Seychelles, Sri Lanka, Uganda, and Ukraine. There is a sharp decline in exchange rate discussions, with an emphasis on inflation, interest rates, and economic activity following the adoption of IT. This shift reflects the framework's focus on anchoring inflation expectations, prompting a transition from backward-looking exchange rate statements to forward-looking discussions on inflation and interest rates. Additionally, references to economic activity become more forward-looking, underscoring the role of output gap assessments in policy decisions. These changes underscore how IT adoption systematically reshapes central bank communication, reinforcing a forward-looking approach in the pursuit of price stability.

We can explicitly measure forward-lookingness in central bank documents by computing the proportion of forward-looking sentences in a document. We call this the *forward-lookingness score*, which is expressed as:

$$\text{Forward-lookingness Score}_{c,d,t} = \frac{\#\text{Forward-Looking}_{c,d,t}}{\#\text{Forward-Looking}_{c,d,t} + \#\text{Backward-Looking}_{c,d,t}}, \qquad (2)$$

in which c, d, t index the central bank/economy, document type (communication outlet), and time.

Figure 14: Forward- and Backward-Looking Monetary Policy Communication in the UK and USA



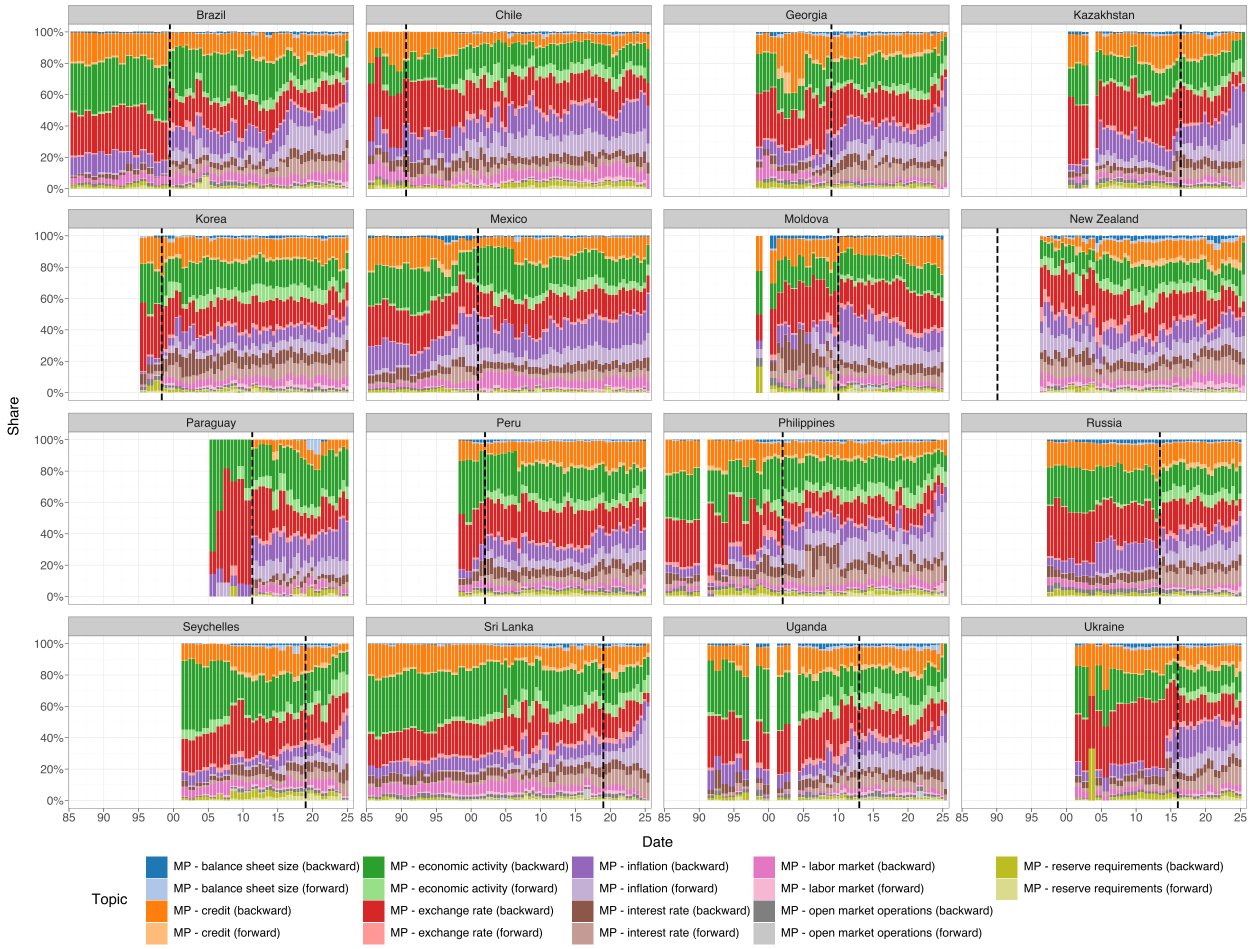
Notes: This figure shows the evolution of monetary policy communication in the United Kingdom and the United States, broken down by monetary policy subtopics and communication stance (forward- and backward-looking). For each document, the number of sentences in each monetary policy subtopic is expressed as a share of total monetary policy sentences. Lighter colors indicate forward-looking content, and darker colors indicate backward-looking content. The vertical dashed line denotes the UK's adoption of inflation targeting, as reported in the IMF's AREAER dataset.

The terms #Forward-Looking and #Backward-Looking represent the number of forward-looking and backward-looking sentences, respectively.

Figure 16a shows that forward-looking communication has increased across most central bank communication outlets. As expected, speeches and monetary policy decisions exhibit the highest forward-looking scores, given their role in shaping expectations and guiding future policy actions. Notably, even traditionally retrospective documents, such as annual reports—which typically assess past economic performance, financial statements, and policy outcomes—increased their forward-looking content. This could be driven by central banks' efforts to enhance transparency by contextualizing past performance within future policy directions, including discussions on prospective autonomy.

One limitation of analyzing the forward-looking score over absolute time is that shifts in the sample composition may influence the observed aggregate trends. To address this, Figure 16b depicts the forward-looking score in terms of elapsed months since each publication type's

Figure 15: Forward- and Backward-Looking Monetary Policy Communication across Inflation-Targeting Economies

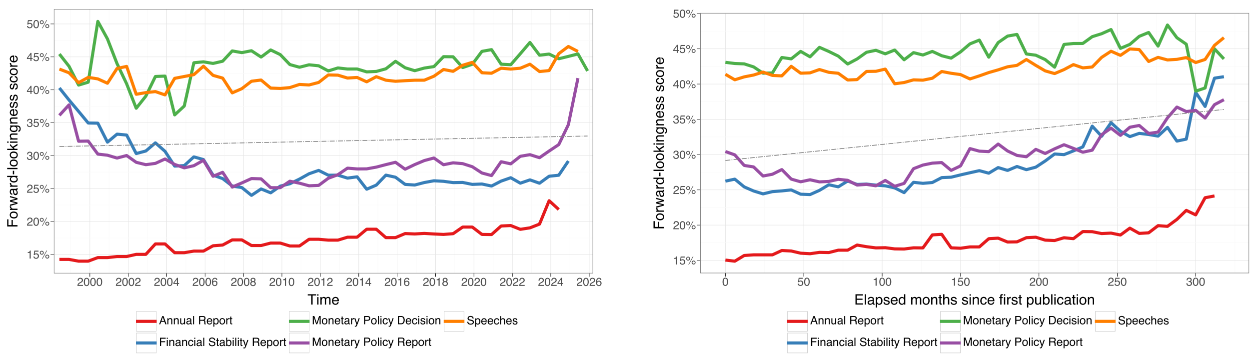


Notes: This figure tracks the evolution of monetary policy communication across selected inflation-targeting economies, broken down by monetary policy subtopics and communication stance (forward- and backward-looking). For each document, the number of sentences in each monetary policy subtopic is expressed as a share of total monetary policy sentences. Lighter colors indicate forward-looking content, and darker colors indicate backward-looking content. The vertical dashed line denotes the economy's adoption of the inflation targeting monetary policy framework, as reported in the IMF's AREAER dataset.

first available document. This perspective confirms that the trend toward more forward-looking communication is systematic rather than an artifact of changing sample composition, including financial stability reports.

To complement the analysis of aggregate averages over time, Figure 17 shows the distribution of forward-lookingness scores globally by communication outlet, overlaying group-specific medians for economies classified by level of development and monetary policy framework. The shaded bands represent the global range (25^{th} to 75^{th} percentiles), offering a benchmark against which the temporal dynamics of individual groups can be evaluated. This perspective reveals that advanced economies consistently exhibit above-median forward-looking communication across all report types. In contrast, pegged economies and low-income countries generally remain below the global

Figure 16: Trends in Forward-Lookingness of Central Bank Communication



(a) By calendar year

(b) By months since first publication

Notes: This figure presents the evolution of forward-lookingness scores in central bank communication, disaggregated by communication outlet. Panel (a) shows the trend over calendar years, while Panel (b) aligns communication outlets by the number of months since their first available publication. Forward-lookingness scores are computed as described in Equation (2). The dashed gray line represents global time trends.

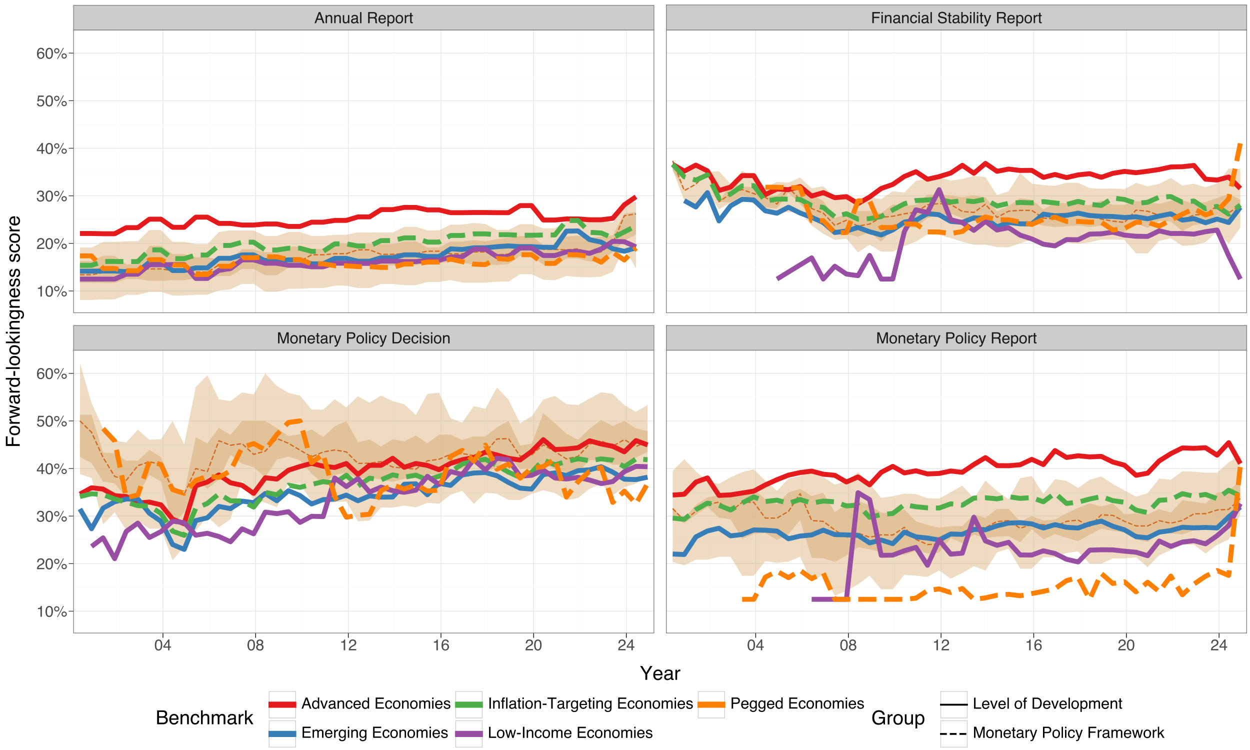
median, particularly in documents in which the forward-looking component is relevant, such as financial stability reports, monetary policy decisions and reports.

The forward-lookingness score, defined in Eq. (2), can be aggregated by document type, topic, or any institutional grouping to examine how central banks adapt their communication strategies over time. Figure 18a shows that forward-looking communication has increased across nearly all central banking traditional or core topics in recent years, notably within monetary policy. This upward trend likely reflects the growing emphasis on expectation management as a core element of the monetary policy transmission mechanism, particularly in inflation-targeting regimes, where effectiveness hinges on shaping agents' expectations of future interest rates and price dynamics. The ability to credibly communicate policy intent reduces informational frictions and enhances the central bank's capacity to steer market outcomes without immediate policy moves.

By contrast, periods of systemic stress, such as the dot-com bust, the Global Financial Crisis, and the COVID-19 pandemic, are associated with a decline in forward-lookingness in crisis-related communication. This reflects a shift toward explaining past shocks and justifying emergency interventions. Meanwhile, communication on emerging structural topics—such as climate change, technological innovation, and structural economic reform, has become markedly more forward-looking (Figure 18b). This trend aligns with the growing integration of these themes into core policy analysis. For instance, climate-related risks are increasingly embedded in monetary policy decisions and financial stability assessments. Similarly, digitalization and structural reforms demand anticipatory frameworks to address evolving systemic challenges. The sustained rise in forward-looking communication in these domains reflects not only the long-term nature of the underlying issues but also the broadening of central bank mandates in response to structural transformations in the global economy.

Figure 19 displays the evolution of the average share of sentences in central bank communication targeted at key audience groups, disaggregated by the level of development. While the financial

Figure 17: Trends in Forward-Lookingness of Central Bank Communication by Level of Development and Monetary Policy Framework

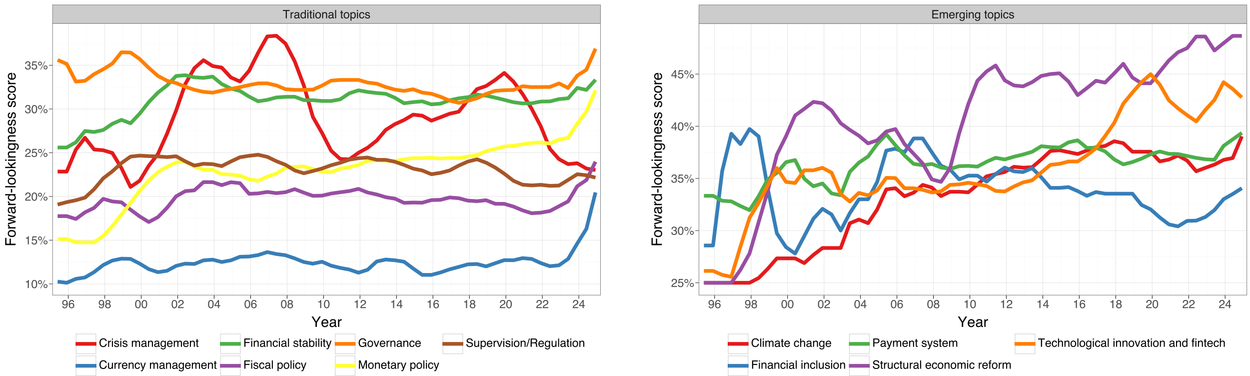


Notes: This figure presents the evolution of forward-lookingness in central bank communications across communication outlets by the economy's level of development (continuous lines) and monetary policy framework (dashed lines). Forward-lookingness scores are computed as described in Equation (2). Shaded bands represent the interquantile ranges (25^{th}–75^{th} percentile) and medians of the forward-lookingness score distribution of all economies by communication outlet.

sector remains the primary audience across all central bank groups, its dominance is declining, most notably in advanced economies. This trend reflects a broadening of central bank outreach beyond traditional market participants, as communication strategies adapt to increasing demands for inclusiveness and accountability. There is an inverse relationship between attention to the government and the level of development. Central banks in low-income and emerging economies systematically direct a larger share of their messaging toward governmental entities. This pattern likely stems from tighter monetary-fiscal linkages, greater reliance on fiscal authorities in macroeconomic stabilization, and the central bank's role as a policy advisor in institutional environments where economic governance is more centralized.

Moreover, the share of communication targeting households has increased across all groups, underscoring a structural shift toward greater public engagement. In advanced economies, the share of communication directed at households and businesses is similar, indicating a balanced approach to outreach to both sectors. By contrast, a notable gap persists in emerging and low-income countries, with businesses receiving more attention than the general public. This disparity may reflect constraints in communication capacity, differences in financial literacy, or a

Figure 18: Forward-Lookingness Trends Across Core and Emerging Central Banking Topics

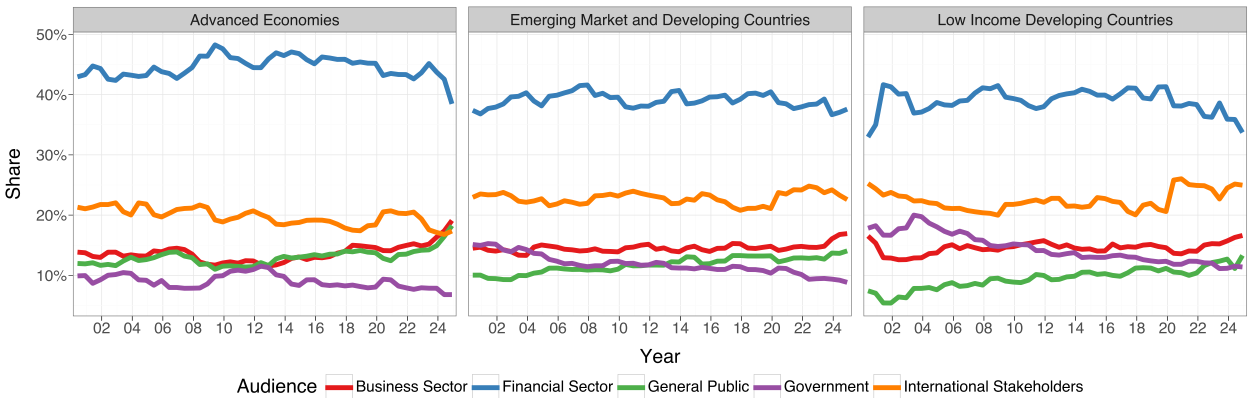


(a) Core central banking topics

(b) Emerging central banking topics

Notes: This figure presents the evolution of forward-lookingness communication across (a) traditional/core and (b) non-traditional/emerging central banking topics. The forward-lookingness score follows Eq. (2).

Figure 19: Audience Targeting in Central Bank Communication by Level of Development



Notes: This figure shows the evolution of the intended audience (main message recipient) in central bank communication, disaggregated by level of economic development. The colors represent different audience categories. Each sentence is classified by its primary target audience. The share of sentences directed at each audience is calculated as a percentage of total sentences in each document. Averages are then taken across all document types published by countries within the same level of development at each point in time.

strategic emphasis on private sector expectations in less developed economies. Overall, the figure indicates that while structural differences in communication priorities persist, a discernible trend toward diversification and tailoring of central bank messaging is evident.

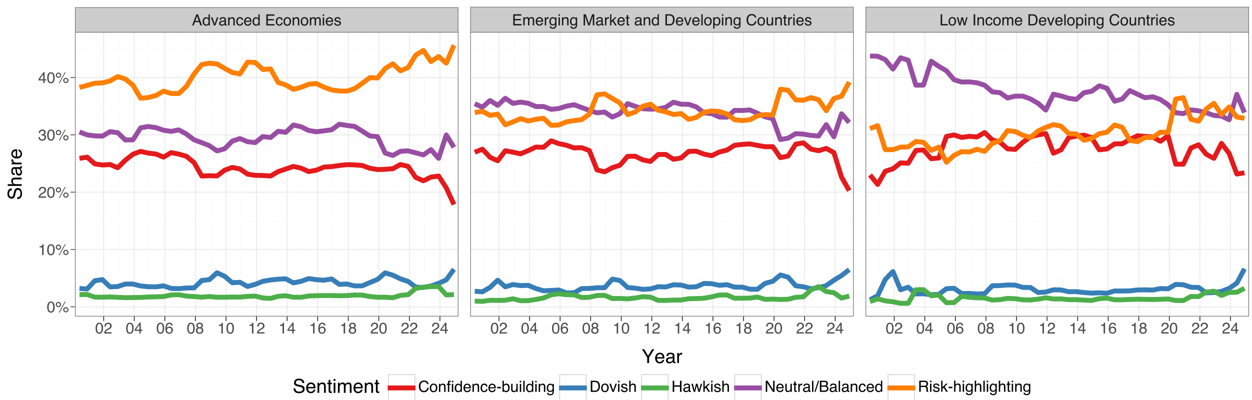
Figure 20 depicts the evolution of sentiment shares in central bank communications, categorized by development level. A striking pattern is the inverse relationship between the prevalence of neutral or balanced content and the level of economic development. Low-income countries often rely heavily on neutral language, characterized by factual reporting and descriptive narratives that provide minimal policy explanation and interpretation. This tendency may reflect limitations in institutional capacity or a deliberate communication strategy tailored to the local audience.

The share of risk-highlighting sentences is broadly proportional to the economy's level of development, suggesting that more mature economies place a higher emphasis on identifying and communicating potential downside risks. This pattern is consistent with the role of risk-aware communication in shaping market expectations in sophisticated financial systems. Risk signaling enables central banks to prepare markets for uncertainty without precommitting to a specific policy stance, thereby maintaining flexibility while anchoring expectations.

Interestingly, the share of confidence-building statements remains relatively stable across all development groups. This consistency implies that, regardless of the level of development, central banks seek to reassure the public, especially during periods of heightened uncertainty. In the broader risk communication architecture, confidence-building is a stabilizing component. At the same time, the relative weight of risk-highlighting varies with the complexity of the financial system and communication objectives.

The observed asymmetry between dovish and hawkish sentiment is also noteworthy. Dovish content appears more frequently than hawkish statements across all development groups. One possible explanation is that accommodative stances are often deployed in response to multifaceted challenges—such as low growth, unemployment, or financial stress—and thus require more elaborate justification. Restrictive policies, by contrast, tend to be justified more succinctly as direct responses to inflationary pressures, particularly in inflation-targeting regimes.

Figure 20: Sentiment Composition of Central Bank Communication by Level of Development



Notes: This figure shows the evolution of sentiment in central bank communication across different levels of economic development. The colors represent different sentiment categories. For each document, the share of sentences expressing a specific sentiment is calculated as a percentage of total sentences. Averages are then computed across all document types published by countries within the same development group at each point in time.

5. Communication Metrics and Their Connection with Financial Variables

This section leverages the sentence-level outputs of the classifier to construct document-level sentiment indicators grounded in economic reasoning. These indicators are then used to assess the predictive content of central bank communication by examining their relationship with market interest rates.

5.1. Methodology

Table 4 summarizes the metrics defined in this section. The net policy sentiment, straightforwardness index, and explanation index use only the monetary policy decisions. In contrast, the net confidence index uses all the regular central bank documents. We discuss their rationale and interpretation below.

Table 4: Definition of Textual Metrics in Central Bank Communications.

Metric	Equation	Description
Net Policy Sentiment (NPS) — *evaluated on monetary policy decisions only*	$NPS = \frac{H-D}{H+D}$	Measures the net stance of monetary policy communication by quantifying the balance between tightening (hawkish) and easing (dovish) signals. Higher values indicate a more restrictive stance, while lower values suggest an accommodative communication.
Straightforwardness Index (SI) — *evaluated on monetary policy decisions only*	$CI = \frac{N+\lvert H-D \rvert}{N+H+D}$	Evaluates the extent to which monetary policy communications convey a dominant policy stance. A lower value suggests the coexistence of conflicting signals or the presentation of multiple policy scenarios, while a higher value indicates clearer and more unidirectional messaging.
Explanation Index (EI) — *evaluated on monetary policy decisions only*	$EI = \frac{C+R+N}{H+D}$	Assesses the level of justification provided in monetary policy decisions. A higher value suggests a more explanatory communication approach, where policy stance statements are supported with contextual information.
Net Confidence Index (NCI) — *evaluated on all regular central bank documents*	$NCI = \frac{C-R}{C+R}$	Captures the central bank's tone by assessing the prevalence of confidence-building versus risk-highlighting statements. A higher index reflects optimism in the economic outlook, while a lower index signals caution.

Notes: H, D, C, R, and N represent the number of hawkish, dovish, confidence-building, risk-highlighting, and neutral statements, respectively.

5.1.1. Net Policy Sentiment

The *Net Policy Sentiment* (NPS) metric quantifies the directional stance of central bank communication by capturing the balance between *hawkish* (tightening) and *dovish* (easing) communication signals. Formally, the NPS for a given document is defined as:

$$NPS = \frac{H - D}{H + D}, \tag{3}$$

in which H and D represent the number of *hawkish* and *dovish* sentences, respectively. The metric ranges from $[-1, 1]$, with positive values indicating a predominance of hawkish communication and negative values reflecting a dovish tone.

We focus on computing this metric within the context of monetary policy decision documents, where the economic interpretation of NPS is most meaningful. In this setting, the NPS captures the *directional stance of communication*, a construct distinct from the actual *monetary policy stance*. While the latter is implemented through instruments such as the policy interest rate or balance sheet operations, the former operates through language, shaping expectations and strengthening monetary policy transmission. In inflation-targeting regimes, the monetary policy decision typically embeds two distinct signals: the *current monetary policy stance* itself, delivered through the quantitative value of the main policy instrument, and the *communication about the future stance*, substantiated through the forward-looking messages in the document. The second type of signal is the *forward guidance*, based on which central banks could provide information about their future monetary policy intentions.

To distinguish between these temporal components, we refine Eq. (3) by disaggregating sentences into forward-looking and backward-looking subsets. We define the *forward-looking* and *backward-looking* NPS components as follows:

$$NPS_{fwd} = \frac{H_{fwd} - D_{fwd}}{H_{fwd} + D_{fwd}}, \tag{4}$$

$$NPS_{bwd} = \frac{H_{bwd} - D_{bwd}}{H_{bwd} + D_{bwd}}, \tag{5}$$

in which subscripts fwd and bwd denote the number of forward-looking and backward-looking sentences, respectively.

These forward- and backward-looking sentiment scores are not directly additive in their raw form. However, the overall NPS can be recovered as a weighted linear combination of the two components:

$$NPS = \omega_{fwd} \cdot NPS_{fwd} + \omega_{bwd} \cdot NPS_{bwd}, \tag{6}$$

where $\omega_{fwd} = \frac{H_{fwd}+D_{fwd}}{H+D}$ and $\omega_{bwd} = \frac{H_{bwd}+D_{bwd}}{H+D}$ represent the relative shares of forward- and backward-looking sentences among all hawkish and dovish sentences in the document.

This decomposition is informative for both theoretical and empirical reasons. Theoretically, the relative weights ω_{fwd} and ω_{bwd} endogenously reflect the central bank's emphasis on forward guidance versus retrospective and current assessments. A higher ω_{fwd} indicates that the communication is more forward-looking, suggesting that the central bank actively uses language to guide future expectations. Conversely, a higher ω_{bwd} reflects a greater focus on explaining past decisions or describing current conditions. This weighted formulation also ensures that changes in the prominence of forward-looking communication are adequately accounted for when evaluating the overall stance.

The *forward-looking NPS* (NPS_{fwd}) serves as a quantifiable proxy for *monetary policy guidance*, offering a direct measure of how central banks employ forward guidance. This approach complements traditional forward guidance measures, which are often based on financial market reactions or survey-based inference (e.g., Gürkaynak et al. (2005)). In contrast, our method derives guidance from textual content, enabling granular tracking of communicative policy shifts over time.

The *backward-looking NPS* (NPS_{bwd}) captures the central bank's retrospective narrative—its emphasis on past economic developments, policy inertia, and assessment of realized outcomes. A central bank displaying a hawkish tone in backward-looking statements while maintaining a dovish forward-looking tone may be signaling that past inflation pressures have subsided, thereby paving the way for a more accommodative policy path.

By disaggregating the net policy sentiment in this way, we obtain a richer representation of monetary policy communication—one that separates ex-post justification from ex-ante guidance and enables systematic study of their respective effects on market expectations and macroeconomic outcomes.

5.1.2. Straightforwardness Index

The *Straightforwardness Index* (SI) systematically measures whether central banks deliver clear and coherent policy signals, assessing the extent to which monetary policy communications convey a unidirectional stance versus presenting multiple potential policy paths. The index is formally defined as:

$$SI = \frac{N + |H - D|}{N + H + D}, \tag{7}$$

where N, H, and D denote the number of neutral, hawkish, and dovish statements in the same monetary policy decision. The numerator aggregates: (i) the absolute net sentiment, $|H-D|$, which reflects the dominance of a particular directional tone; and (ii) the number of neutral statements, which contribute contextual clarity without signaling direction. The denominator normalizes by the total number of policy-relevant statements, ensuring comparability across documents of varying length and scope.

By construction, $SI \in [0, 1]$, with higher values indicating a more internally consistent and decisive communication. An SI approaching 1 implies a clear dominance of one sentiment category—either hawkish or dovish—supported or not by context-setting neutral statements. In contrast, values closer to 0 indicate that hawkish and dovish elements coexist similarly, thereby diluting the dominant signal.

While the aggregate SI provides an overall measure of clarity, further insight can be gained by decomposing the index according to the communication stance of the underlying statements. Specifically, we calculate separate indices for forward- and backward-looking sentences:

$$SI^{(s)} = \frac{N^{(s)} + |H^{(s)} - D^{(s)}|}{N^{(s)} + H^{(s)} + D^{(s)}}, \quad s \in \{\text{Forward}, \text{Backward}\}, \tag{8}$$

where s indexes the stance of the sentence, and $N^{(s)}$, $H^{(s)}$, and $D^{(s)}$ refer to neutral, hawkish, and dovish statements within that stance category. This decomposition recognizes that clarity serves different roles in retrospective and prospective communication.

Backward-looking communication typically involves factual reporting and justification of past policy decisions. As such, it is generally more straightforward, reflecting observed outcomes and unambiguous rationales. In contrast, forward-looking communication necessarily incorporates uncertainty and conditionality. Statements about the future often outline alternative scenarios and policy paths, which inherently reduce straightforwardness. A lower $SI^{(\text{Forward})}$ is therefore not necessarily undesirable: it may indicate a deliberate strategy of conveying contingency and flexibility in response to evolving economic conditions.

This decomposition provides important analytical leverage. Comparing $SI^{(\text{Forward})}$ and $SI^{(\text{Backward})}$ allows us to distinguish between communications that are ambiguous due to internal inconsistency (low backward-looking SI) and those that are nuanced and state-contingent (low forward-looking SI). Moreover, variation across country groups in $SI^{(\text{Forward})}$ may reflect differences in institutional capacity. In advanced economies, lower forward-looking straightforwardness is often a feature of sophisticated communication strategies emphasizing conditional guidance. By contrast, in emerging and low-income countries, very low values may also signal limited capacity to articulate clear future guidance or underlying uncertainty in policymaking itself. In this way, the forward- and backward-looking decomposition enhances the interpretability of the index and its ability to reveal underlying features of the communication framework.

5.1.3. Explanation Index

The *Explanation Index* (EI) quantifies the extent to which sentences in monetary policy decisions justify policy actions. It measures the presence of explanatory content—such as discussions of economic conditions, risks, and expressions of confidence—relative to the number of directional statements in a monetary policy decision. Formally, the index is defined as:

$$EI = \frac{C + R + N}{H + D}, \tag{9}$$

in which H and D represent the number of hawkish and dovish statements, respectively, while C, R, and N denote the number of confidence-building, risk-highlighting, and neutral statements. The numerator captures the volume of content typically associated with justification and policy narrative, while the denominator captures sentences conveying a clear policy direction.

The EI reflects how actively a central bank engages in *justifying its actions*. A higher value implies that directional signals are embedded within a broader communicative framework, explaining the rationale behind decisions, elaborating on trade-offs, or contextualizing the policy path. This is particularly important in settings of heightened economic uncertainty or when a central bank seeks to build transparency. Conversely, a lower EI may indicate a more concise communication strategy. Importantly, a low EI does not necessarily imply poor communication. Lower values may also reflect deliberate restraint, especially during elevated uncertainty when central banks avoid issuing overly specific guidance that could later prove misguided. For instance, during the onset of the COVID-19 pandemic, many advanced economy central banks reduced explanatory language to preserve flexibility and avoid sending potentially misleading signals. More generally, in well-established monetary frameworks with strong reputations, concise statements may suffice to anchor expectations. However, in environments where transparency is imperfect or markets demand more explanation, limited explanatory content may increase uncertainty and weaken the effectiveness of policy signaling.

The explanation index offers a structured, quantitative alternative to traditional proxies for explanatory richness. Earlier approaches have relied on qualitative assessments of policy reports (e.g., Blinder et al. (2008)) or used document length as a proxy for explanation depth (e.g., Hansen et al. (2017)), assuming that longer statements reflect greater justification. Yet, verbosity does not equate to clarity; extended texts may include repetition or generic language without substantive content. Unlike readability scores that focus on linguistic complexity (Loughran & McDonald, 2011), the explanation index differentiates between the functional roles of sentences, distinguishing between statements that signal policy stance and those that support, explain, or contextualize that stance.

5.1.4. Net Confidence Index

The *Net Confidence Index* (NCI) captures the tone of central bank communication along the risk-perception spectrum. It measures the balance between statements that reinforce confidence in economic and financial stability and those that highlight potential risks or vulnerabilities. Formally, the index is defined as:

$$NCI = \frac{C - R}{C + R}, \tag{10}$$

where C and R represent the number of confidence-building and risk-highlighting statements, respectively. By construction, the index ranges from -1 (exclusively risk-oriented) to 1 (exclusively confidence-building), with values closer to zero indicating a more balanced or neutral tone.

To account for the temporal orientation of central bank communication, we further decompose the index into forward- and backward-looking components. This distinction allows us to assess whether central banks express confidence or concern about prospective developments or retrospective conditions. The forward-looking and backward-looking indices are defined as:

$$NCI^{(s)} = \frac{C^{(s)} - R^{(s)}}{C^{(s)} + R^{(s)}}, \quad s \in \{\text{Forward}, \text{Backward}\}, \tag{11}$$

where $C^{(s)}$ and $R^{(s)}$ represent the number of confidence-building and risk-highlighting statements classified as forward- or backward-looking, respectively. The forward-looking component, $NCI^{(\text{Forward})}$, reflects the central bank's expectations about future risks and resilience, serving as a proxy for the institution's risk outlook. In contrast, the backward-looking component, $NCI^{(\text{Backward})}$, captures retrospective assessments of prevailing or realized conditions.

Similar to the other indices, the overall NCI can be expressed as a weighted average of its forward- and backward-looking components:

$$NCI = \omega^{(\text{Forward})} \cdot NCI^{(\text{Forward})} + \omega^{(\text{Backward})} \cdot NCI^{(\text{Backward})}, \tag{12}$$

where $\omega^{(s)} = \frac{C^{(s)}+R^{(s)}}{C+R}$ denotes the share of statements within each temporal orientation. These weights ensure that the relative prominence of future- and past-oriented communication is appropriately reflected in the aggregate index.

Decomposing the NCI yields valuable insights. The gap between forward- and backward-looking components captures shifts in the tone of risk communication. A positive gap (i.e., $NCI^{(\text{Forward})} > NCI^{(\text{Backward})}$) suggests improving sentiment, with the central bank expressing greater confidence about future prospects than about current or past conditions. Conversely, a negative gap may signal deteriorating expectations, with forward-looking statements becoming more risk-focused relative to retrospective assessments, potentially serving as an early indicator of heightened uncertainty.

The decomposition also clarifies the dynamic roles of risk communication. Backward-looking statements tend to rationalize or contextualize recent developments and thus align closely with contemporaneous indicators of financial conditions, such as the VIX. Forward-looking statements, by contrast, convey anticipatory signals about potential risks and resilience, contributing to expectation formation and potentially affecting market volatility before shocks materialize.

The interpretation of the NCI varies by communication outlet. In monetary policy decisions, it primarily captures macro-financial concerns such as inflation risks. When derived from financial stability reports, it reflects the central bank's assessment of systemic vulnerabilities and resilience. In broader documents—such as annual reports and speeches—the index conveys a composite view of institutional confidence and perceived risks across monetary, financial, and structural policy areas.

5.2. Empirical Application on the Central Bank Communications Data

This section applies the metrics discussed before to the global central bank communications dataset.

5.2.1. Net Policy Sentiment

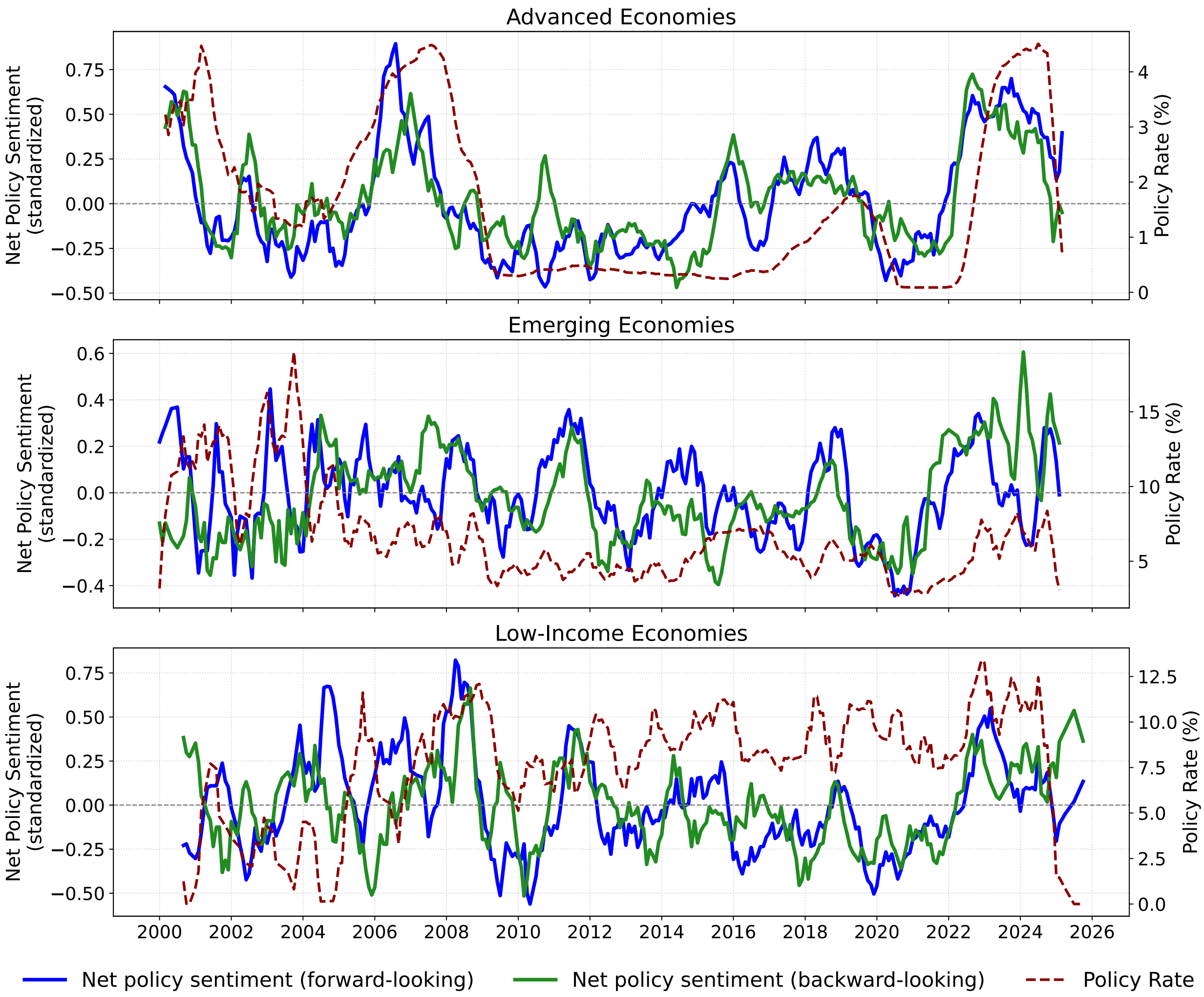
Figure 21 traces the evolution of net policy sentiment—disaggregated into forward- and backward-looking components—across advanced, emerging, and low-income economies. Each series is weighted by country-level nominal GDP (in U.S. dollars), ensuring that the global index reflects the tone of systemically important monetary authorities. The curves are standardized for each group of economies. This transformation expresses deviations from each group's historical average in standard deviation units, allowing us to abstract from structural differences in average tone—such as chronically hawkish or dovish biases—and focus on shifts in the underlying communication stance over time.

Three key findings emerge. First, the forward-looking component of net policy sentiment frequently anticipates changes in policy rates, especially in advanced economies. This leading behavior underscores the forward-looking content carried by the net policy sentiment. It is consistent with the role of central bank communication in shaping expectations, especially in inflation-targeting monetary policy frameworks. A tightening in the forward-looking net policy sentiment typically precedes actual rate increases, indicating that the index captures information relevant to future policy moves, beyond contemporaneous economic conditions.

Second, the forward-looking net policy sentiment embeds elements of monetary policy communication beyond policy rate changes. This feature becomes evident during episodes where conventional policy tools were constrained. After the global financial crisis and during the COVID-19 pandemic, central banks in advanced economies operated near the effective lower bound of interest rates. While policy rates remained flat, the net policy sentiment varied substantially, reflecting directional stance textual signals embodied in other monetary policy tools. These included elements of forward guidance, balance sheet and asset purchase programs, as well as other forms of unconventional monetary policy tools. Thus, the net policy sentiment serves as a comprehensive proxy for the effective monetary policy stance.

Third, the alignment between net policy sentiment and realized policy rates varies substantially across development levels. Forward-looking sentiment closely aligns with policy rates in advanced economies, underscoring the consistent use of communication as a policy instrument in these economies. In contrast, the relationship is weaker in emerging and low-income economies. This decoupling likely reflects a combination of structural and institutional constraints, including limited

Figure 21: Net Policy Sentiment and Policy Rates by Level of Development

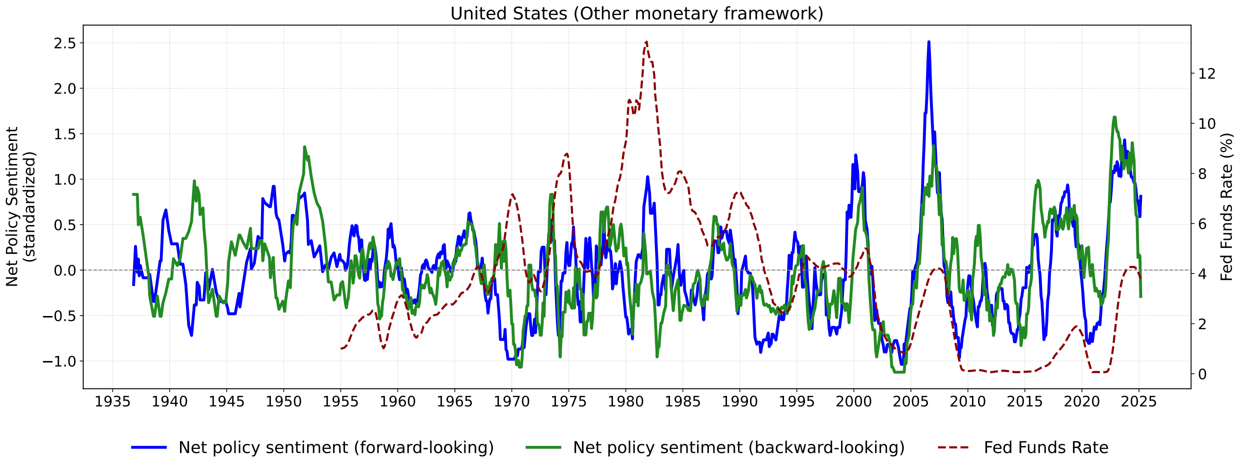


Notes: This figure compares average net policy sentiment (left vertical axis) and short-term policy rates (right vertical axis) across economies grouped by level of development. Net policy sentiment is decomposed into forward- and backward-looking components and standardized within each group to remove structural differences in tone and facilitate comparison. Policy rates are sourced from the IMF's International Financial Statistics (IFS). All variables are weighted by country-level nominal GDP in US dollars to emphasize systemically important economies. The horizontal dashed line indicates each group's long-run average net policy sentiment, serving as a benchmark for neutrality.

use of forward guidance, weaker policy transmission mechanisms, and greater vulnerability to credibility shocks in less advanced economies.

These results highlight the potential of text-based indicators to reconstruct the effective monetary policy stance, even in environments where data on the policy rate (or any other proxy) are limited. In the case of the United States, the Federal Reserve has published monetary policy decisions since 1936, allowing us to extend the forward-looking net policy sentiment before the actual Fed Funds rate series that started in 1954. As shown in Figure 22, the sentiment-based

Figure 22: Net Policy Sentiment and Policy Rate in the United States



Notes: This figure compares the net policy sentiment index (left vertical axis) and the federal funds rate (right vertical axis) for the United States. The sentiment index is decomposed into forward-looking and backward-looking components. The policy rate is sourced from the Federal Reserve Economic Data (FRED). The sentiment index is standardized relative to the historical distribution of the United States, accounting for structural tone shifts and allowing interpretation of deviations from the country's long-run communication norm. The horizontal dashed line represents the U.S. historical average net policy sentiment, serving as a neutral benchmark.

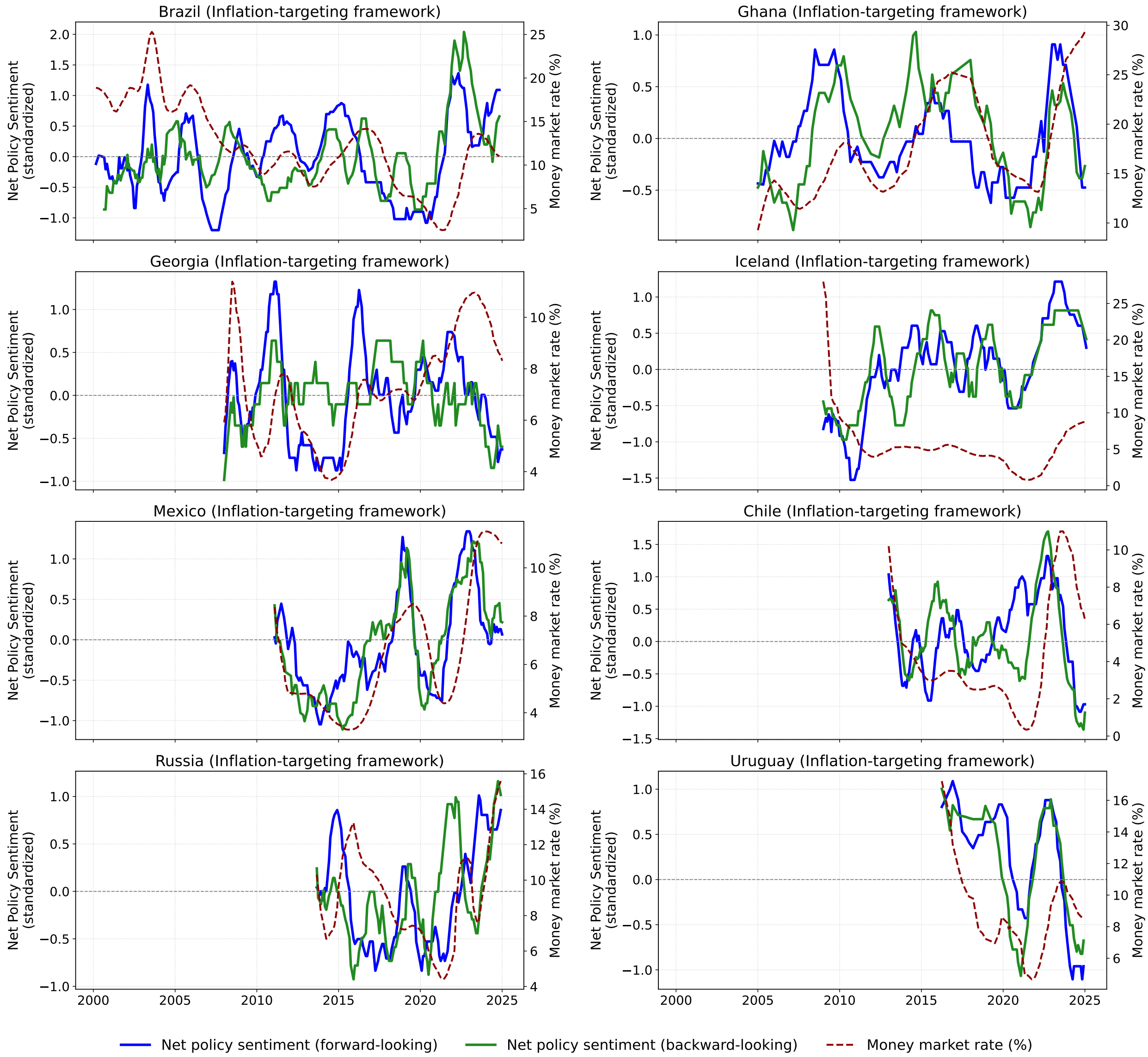
indicator roughly tracks the federal funds rate once it becomes available, particularly in more recent periods such as the 2001 and 2008 recessions, and the COVID-19 tightening cycle.

This approach is particularly valuable for assessing whether central bank communication aligns with the intended monetary policy stance—an issue especially relevant in low-income and some emerging economies. In these contexts, policy frameworks are often less transparent, and implementation capacity may be limited. Standardized policy rate data are frequently unavailable or discontinuous, further complicating efforts to track the policy stance over time. By relying on textual indicators extracted from policy communications, our tool provides a systematic way to monitor and assess whether central bank signals align with observed short-term interest rates.[13]

Figure 23 compares forward- and backward-looking components of net policy sentiment with interbank money market rates across a sample of inflation-targeting emerging economies. Overall, the net policy sentiment—particularly its forward-looking component—closely tracks money market rates in most inflation-targeting economies. In some cases, however, we observe delayed or muted responses, with the money market rate adjusting more gradually to changes in tone. The responsiveness of the money market rate is shaped by structural and institutional frictions. These include liquidity management, the intention to signal a policy stance without incurring the associated costs for the central bank, and less liquid interbank markets. Additionally, central bank credibility can influence how quickly policy signals are translated into market

[13] Due to limited availability of policy rate data for these economies in the IMF's IFS dataset, we rely on money market rates as an alternative. While the policy rate reflects the announced target, money market rates capture actual liquidity conditions and implementation. In economies with underdeveloped interbank markets or weak transmission mechanisms, the two can diverge substantially.

Figure 23: Net Policy Sentiment and Money Market Rates in Selected Economies



Notes: This figure compares the net policy sentiment (left vertical axis) and short-term interbank money market rates (right vertical axis) for selected economies with inflation-targeting frameworks. The sentiment index is decomposed into forward-looking and backward-looking components. Money market rates are sourced from the IMF's International Financial Statistics (IFS). Sentiment indices are standardized relative to each country's historical distribution to adjust for structural tone differences and highlight deviations from country-specific communication norms. The horizontal dashed line represents each country's historical average net policy sentiment, serving as a neutral benchmark.

expectations and interbank rates. These factors suggest that, while communication plays a crucial role in shaping monetary conditions across all economies, the speed and effectiveness of its transmission vary significantly depending on the level of financial market development and institutional characteristics.

We now turn to a formal panel-data analysis to evaluate whether the net policy sentiment measure, particularly its forward-looking component, effectively captures the stance of monetary policy communicated by central banks. Data is monthly. Table 1 provides summary statistics of all variables used in this section and details about their construction. This exercise is not designed to establish causality. Instead, it aims to test the internal consistency between a central bank's stated narrative and its actual policy stance on a global scale using our central bank communication dataset. If the sentiment measure reflects genuine policy intent, we would expect a more hawkish (or dovish) tone to precede or coincide with higher (or lower) policy rates. We use the following specification:[14]

$$\text{Policy Rate}_{i,t} = \alpha_i + \lambda_t + \eta \text{ Policy Rate}_{i,t} + \beta \, NPS_{i,t} + X'_{i,t}\gamma + \varepsilon_{i,t} \tag{13}$$

in which i and t index the country and time, respectively. The dependent variable, Policy Rate$_{i,t}$, denotes the policy rate of economy i at time t. The main regressor, NPS$_{i,t}$, includes either the total, forward-looking, or backward-looking component of net policy sentiment. The control vector $X_{i,t}$ comprises macroeconomic fundamentals, including consumer price inflation and the exchange rate, as well as communication-related metrics: the explanation index, straightforwardness index, and net confidence index. We also include the first lagged policy rate to account for monetary policy inertia. Conceptually, adding the lagged dependent variable permits us to interpret the contribution of the remaining variables to explain variations in the policy rate from $t-1$ to t. $\varepsilon_{i,t}$ represents the error term, capturing idiosyncratic shocks not accounted for by the included regressors. The dependent and independent variables are standardized by country, allowing for interpretation in terms of standard deviations from each country's historical average. We cluster errors at the country level to account for serial correlation within each country across time.

The term α_i represents country fixed effects and accounts for time-invariant country-specific features, thereby mitigating concerns about omitted variable bias arising from non-observed, invariant structural differences among economies. These fixed effects capture persistent (or roughly persistent) institutional characteristics such as central bank independence, monetary policy frameworks, and historical credibility in monetary policymaking—factors that could systematically influence communication strategies and policy rate decisions. The term λ_t embodies time-specific fixed effects and absorbs global macroeconomic shocks that affect all countries simultaneously. Given the long period of our panel, this is crucial to controlling for external conditions such as financial crises, commodity price shocks, and shifts in the global interest rate environment.

Table 5 presents coefficient estimates from four specifications based on Equation (13), each probing the relationship between central bank communication and the prevailing policy rate. Odd-numbered specifications use the total net policy sentiment, and even-numbered specifications use the forward- and backward-looking components as covariates. Specifications (I) and (II) use the full sample. Specifications (III) and (IV) replicate these models using a restricted subsample of countries with at least 10 years of data, helping to mitigate dynamic panel bias in settings with

[14] The net policy sentiment data is recorded as of the monetary policy decision date, which falls on or before the last day of each month. Similarly, the policy rate series is constructed using end-of-month values. Therefore, the timing convention ensures that our regressions do not suffer from lookahead bias, as sentiment indicators are observed prior (in the same month) to or contemporaneously with the policy rate data.

shorter time series (Nickell, 1981).[15]

Across all specifications, net policy sentiment exhibits a statistically significant and economically meaningful association with the prevailing policy rate. Focusing on Specification (I), the coefficient on total net policy sentiment is 0.031. Given that the sample standard deviation of the policy rate is 6.36 percentage points (Table 1 in Appendix E), this implies that a one-standard-deviation increase in the total net policy sentiment is associated with an increase of approximately 20 basis points ($0.031 \times 6.36 \approx 0.20$ p.p.) in the policy rate. Because the model includes a lagged dependent variable, this effect reflects the adjustment from one monetary policy decision to the next, conditional on the existing rate. The economic relevance of this result is underscored when compared to the actual distribution of interest rate changes in the sample: the absolute policy rate change is 25 basis points or lower for 75 percent of all policy rate changes in absolute terms (Table 1 in Appendix E). In this context, a 20-basis-point shift attributable to sentiment alone represents a substantial share of observed rate movements. When decomposed in Specification (II), both the forward- and backward-looking components remain significant. However, the forward-looking coefficient is relatively larger, consistent with the role of anticipatory communication in conveying policy direction.

Taken together, these results support the internal consistency of central bank communication: the language used in monetary policy decisions is reflected in policy rates. These findings align with theoretical models that emphasize communication as a policy tool (Woodford, 2005) and empirical evidence on the signaling role of forward guidance (Campbell et al., 2012). Although the analysis is not designed to establish causality, it highlights that the net policy sentiment derived from textual information in monetary policy decisions provides a clear signal of the effective monetary stance at the global scale using our unique dataset.

[15] In dynamic panel models of the form $y_{i,t} = \alpha_i + \lambda_t + \rho y_{i,t-1} + X_{i,t}\beta + \varepsilon_{i,t}$, the inclusion of the lagged dependent variable induces correlation with the fixed effects, resulting in a downward bias known as the Nickell bias (Nickell, 1981). While GMM-based corrections are available (Ahn & Schmidt, 1995), the bias decreases with longer time dimensions. Given our monthly panel and long time span, this concern is less acute in our setting.

Table 5: Communication consistency test: is the prevailing policy stance reflected in central bank communication?

	Policy Rate$_{i,t}$			
	(I)	(II)	(III)	(IV)
Net Policy Sentiment$_{i,t}$	0.031***		0.028***	
	(0.005)		(0.006)	
Net Policy Sentiment (Forward)$_{i,t}$		0.026***		0.021***
		(0.004)		(0.005)
Net Policy Sentiment (Backward)$_{i,t}$		0.015***		0.012***
		(0.004)		(0.004)
Policy Rate$_{i,t-1}$	0.970***	0.968***	0.975***	0.973***
	(0.005)	(0.005)	(0.005)	(0.005)
Net Confidence Index$_{i,t}$	-0.002	-0.003	-0.002	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)
Explanation Index$_{i,t}$	0.002	0.001	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)
Decisiveness Index$_{i,t}$	0.002	0.003	0.004	0.006
	(0.004)	(0.004)	(0.004)	(0.004)
Inflation (CPI)$_{i,t}$	0.000	0.003	0.015	0.019*
	(0.012)	(0.011)	(0.012)	(0.010)
Exchange Rate (USD/local)$_{i,t}$	-0.008	-0.008	-0.009**	-0.010***
	(0.005)	(0.005)	(0.004)	(0.004)
Country Fixed Effects	x	x	x	x
Time Fixed Effects	x	x	x	x
Sample	Full	Full	≥ 10 years	≥ 10 years
Observations	5209	5067	3913	3807
R^2	0.970	0.969	0.980	0.980

Notes: This table reports coefficient estimates from panel fixed-effects regressions where the dependent variable is the policy rate of country i at time t. The main explanatory variables are net policy sentiment metrics derived from monetary policy decisions, including total sentiment and its forward- and backward-looking components. Additional controls include textual communication indices (net confidence, explanation, and decisiveness), inflation (CPI), exchange rate (USD/local), and the lagged policy rate. All variables are standardized by country. Specifications (III)–(IV) restrict the sample to countries with at least 10 years of data. Country and time fixed effects are included in all regressions. Standard errors are clustered at the country level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

To assess whether the communication–policy rate link varies across institutional settings, Table 6 reports regressions stratified by monetary policy framework. The net policy sentiment is most strongly associated with policy rates in inflation-targeting regimes, where both total and forward-looking components are highly significant. This aligns with theory: in discretionary frameworks reliant on managing expectations, communication serves as an important monetary policy tool, especially the forward-looking component. By contrast, no significant relationship emerges under exchange rate anchor regimes, where the policy rate is typically subordinated to the exchange rate objective, reducing the scope for communication to influence domestic rate setting.

Results for "other frameworks" are mixed, reflecting the heterogeneity of regimes with hybrid or evolving mandates. These findings underscore that the signaling power of communication is conditional on institutional context.

Table 6: Heterogeneity in the communication–policy link across monetary frameworks

	Policy $\text{Rate}_{i,t}$							
	Inflation-targeting (I)	Monetary aggregate (II)	Other framework (III)	Exchange rate anchor (IV)	Inflation-targeting (V)	Monetary aggregate (VI)	Other framework (VII)	Exchange rate anchor (VIII)
Net Policy $\text{Sentiment}_{i,t}$								
Total	0.029***	0.057	0.036**	-0.068				
	(0.006)	(0.029)	(0.015)	(0.065)				
Forward-looking					0.024***	0.080**	0.035*	0.014
					(0.005)	(0.029)	(0.015)	(0.036)
Backward					0.014***	0.014	0.022***	-0.057
					(0.004)	(0.012)	(0.005)	(0.053)
Straightforward. $\text{Index}_{i,t}$	0.004	-0.022	0.004	0.106	0.005	-0.038*	-0.001	0.089
	(0.004)	(0.016)	(0.012)	(0.065)	(0.004)	(0.015)	(0.010)	(0.057)
Explanation $\text{Index}_{i,t}$	0.004	0.019	0.008	0.011	0.003	0.022	0.012	0.050
	(0.003)	(0.010)	(0.016)	(0.031)	(0.003)	(0.012)	(0.013)	(0.027)
Net Confidence $\text{Index}_{i,t}$	-0.006*	-0.015	0.012*	-0.056	-0.007**	-0.020	0.014***	-0.067
	(0.003)	(0.014)	(0.005)	(0.038)	(0.003)	(0.014)	(0.004)	(0.047)
Inflation $(\text{CPI})_{i,t}$	0.010	0.076	-0.019	-0.212**	0.011	0.078	-0.018	-0.267***
	(0.012)	(0.056)	(0.023)	(0.053)	(0.012)	(0.054)	(0.025)	(0.042)
Exchange $\text{Rate}_{i,t}$	-0.012**	0.051	-0.025	0.140*	-0.012**	0.054	-0.025	0.123*
	(0.005)	(0.048)	(0.031)	(0.045)	(0.005)	(0.043)	(0.031)	(0.043)
Policy $\text{Rate}_{i,t-1}$	0.965***	1.030***	0.952***	0.874***	0.963***	1.022***	0.949***	0.838***
	(0.005)	(0.020)	(0.006)	(0.021)	(0.006)	(0.021)	(0.007)	(0.014)
Country Fixed Effects	x	x	x	x	x	x	x	x
Time Fixed Effects	x	x	x	x	x	x	x	x
Observations	4164	238	580	227	4034	237	571	225
R^2	0.974	0.985	0.970	0.985	0.974	0.985	0.970	0.986

Notes: This table reports fixed-effects panel regression estimates where the dependent variable is the standardized policy rate. Specs. (I)–(IV) use the total Net Policy $\text{Sentiment}_{i,t-1}$ derived from monetary policy decisions. Specs. (V)–(VIII) decompose the backward- and forward-looking components of the net confidence index. All specifications include controls for communication indices (net confidence index, explanation index, straightforwardness index), macroeconomic conditions (inflation, exchange rate – USD/local), and the lagged dependent variable. We include country and time fixed effects in all specifications. All variables are standardized at the country level. Standard errors are clustered by country. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

While the previous empirical exercise established that net policy sentiment aligns with contemporaneous policy decisions, suggesting internal coherence in communication, it remains unclear whether such communication also contains forward-looking informational content relevant to subsequent policy and market developments. To test this, we estimate panel-data regressions that examine whether lagged sentiment indicators are systematically associated with future changes in interest rates, controlling for macroeconomic variables and fixed effects. This analysis shifts the focus from contemporaneous coherence between words and actions to whether the tone of communication at time t is systematically linked to policy and market outcomes at time $t + 1$, conditional on fundamentals and institutional factors. As monetary policy increasingly operates through expectations and forward guidance, this analysis provides a direct test of how effectively central banks use communication to shape the future trajectory of interest rates across different maturities.

The estimated equation takes the form:

$$\Delta \text{Rate}_{i,t+1} = \alpha_i + \lambda_t + \beta \text{ Net Policy Sentiment}_{i,t} + X'_{i,t}\gamma + \varepsilon i, t, \tag{14}$$

where $\Delta\text{Rate}_{i,t+1}$ denotes the one-period-ahead change in either the policy rate, the short-term (T-bill) rate, or the long-term (T-bond) rate. The variable of interest, Net Policy Sentiment$_{i,t}$, is added in total terms or segmented into forward- and backward-looking components. While this decomposition adds interpretability, forward- and backward-looking components often co-move, especially during turning points, raising the question of whether markets respond to the level of forward-looking sentiment or its prominence relative to retrospective narratives. To address this, we augment the analysis with an alternative specification that introduces the *gap* between forward- and backward-looking sentiment, alongside the forward-looking component, as explanatory variables. This allows us to test whether markets react not only to the absolute tone of forward-looking communication but also to shifts in the communicative balance towards signaling future policy intentions. The fixed effects and macroeconomic controls (lagged one period) mirror those in the consistency specification.

Table 7 presents coefficient estimates from Eq. (14), assessing whether central bank communication aligns with future realized changes in interest rates. The three panels correspond to different rate types: changes in the policy rate (Specs. I–III), short-term treasury bill rates (Specs. IV–VI), and long-term bond interest rates (Specs. VII–IX). Across specifications, forward-looking sentiment is consistently positive and statistically significant, indicating that central bank communication conveys meaningful information about future policy direction and shapes market expectations. Notably, this association remains robust even when controlling for macroeconomic variables such as inflation and exchange rates. Moreover, the gap between forward- and backward-looking sentiment is statistically significant for policy rates, suggesting that markets interpret increases in forward-looking emphasis as especially powerful signals about the policy path ahead.

While central banks set the policy rate directly, market-determined interest rates are sensitive to expectations of future policy actions and thus provide a natural setting to evaluate the informational content of central bank communication. The consistent forward association between sentiment and interest rate changes provides empirical evidence supporting the signaling role of central bank communication. In contrast, the absence of a similar association for backward-looking sentiment suggests that markets differentiate between narrative elements that convey future intentions and those that rationalize past decisions. The muted effect on long-term bonds is also intuitive, as these rates reflect broader factors—including term premia, inflation expectations, and fiscal conditions—that dilute the marginal effect of short-term policy signals. Taken together, the results highlight the role of forward-looking communication as a monetary policy tool.

It is important to clarify the scope of this analysis. The primary objective here is not to establish causality or to generate forecasts, but rather to assess whether communication contains forward-looking information that is systematically linked to policy and market outcomes. In this context, the key result is the robust statistical significance of sentiment even after controlling for macroeconomic fundamentals, exchange rates, and fixed effects. The R^2 values should be interpreted accordingly. In emerging and low-income economies, explanatory power is moderate—generally between 10 percent and 20 percent—reflecting the inherent difficulty of explaining interest rate adjustments in more volatile and policy-diverse environments. In advanced economies, by contrast, the R^2 rises substantially, often exceeding 50 percent, likely due to clearer communication frameworks and closer market attention. Finally, the magnitude of the estimated coefficients reinforces the economic significance of the results.[16]

We extend our previous analysis by evaluating the persistence of the explanatory power of central bank communication across different horizons. To this end, we re-estimate Eq. (14) across multiple leads, regressing changes in policy rates, T-bill rates, and T-bond rates separately at each horizon from 1 to 10 policy decisions ahead. This enables us to track the dynamic alignment between the sentiment embedded in central bank communications and subsequent interest rate movements over time.

Figure 24 presents the estimated coefficients for forward- and backward-looking net policy sentiment, along with 95 percent confidence intervals. The results reveal three core findings. First, forward-looking sentiment maintains a statistically significant association with future changes in the policy rate up to a further horizon, with peak magnitude in the early horizons. This confirms that policy intent communicated at time t is systematically related to the path of monetary policy in the near to medium term. Second, forward-looking sentiment is also significantly associated with changes in short-term market rates (T-bills), particularly over the first five leads. This establishes empirical evidence that markets internalize and act upon the anticipatory content of central bank communication. Third, forward-looking sentiment loses statistical significance when explaining changes in long-term bond interest rates beyond the first lead. This is consistent with the fact that long-term interest rates embed broader macroeconomic expectations, which dilute the signal contained in near-term forward guidance. In contrast, backward-looking sentiment fails to explain future movements in either short-term or long-term market rates, exhibiting weak associations with the policy rate at very short horizons.

We next investigate whether the effect of central bank communication varies systematically

[16] Given that both the dependent variables and sentiment measures are standardized, the coefficients can be interpreted as the change in the dependent variable (in standard deviations) associated with a one standard deviation increase in forward-looking net policy sentiment. Translating these effects into original units indicates that the impact is non-negligible. For example, the coefficient of 0.083 for policy rate changes implies that a one standard deviation increase in sentiment is associated with a future policy rate change of approximately 7 basis points, given the sample standard deviation of policy rate changes (0.87 percentage points, see Table 1 in Appendix E). This magnitude corresponds to 8 percent of the observed median policy rate. Similarly, for short-term market rates (T-bill rates), the coefficient of 0.062 corresponds to a movement of roughly 8 basis points, based on the sample standard deviation of 1.25 percentage points.

across economies at different levels of development. Table 8 reports coefficient estimates from regressions analogous to those in Table 7, but disaggregated by the level of development. We compare advanced economies with the category "other levels of development," encompassing emerging and low-income countries. All variables are standardized within each subsample to ensure comparability of coefficient magnitudes within the same groups.

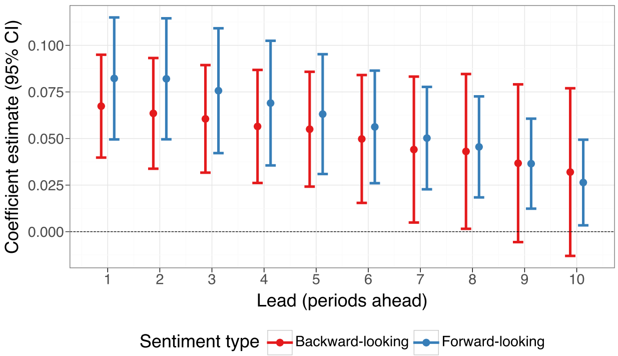
Table 7: Do forward-looking communications align with future policy and market interest rates?

	Δ Policy Rate$_{i,t+1}$			Δ T-Bill Rate$_{i,t+1}$			Δ T-Bond Rate$_{i,t+1}$		
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)
Net Policy Sentiment$_{i,t}$									
Total	0.101***			0.022			0.102*		
	(0.022)			(0.021)			(0.057)		
Forward-looking		0.083***	0.162***		0.055***	0.051***		0.105*	0.135*
		(0.017)	(0.033)		(0.020)	(0.012)		(0.058)	(0.070)
Backward-looking		0.062***			-0.025			0.024	
		(0.014)			(0.019)			(0.015)	
Gap (Fwd – Bwd)			-0.096***			0.038			-0.036
			(0.022)			(0.029)			(0.024)
Straightforward. Index$_{i,t}$	-0.010	0.006	0.006	0.021	0.023	0.023	-0.022	-0.006	-0.006
	(0.011)	(0.009)	(0.009)	(0.021)	(0.018)	(0.018)	(0.028)	(0.019)	(0.019)
Explanation Index$_{i,t}$	-0.003	-0.006	-0.006	-0.010	-0.013	-0.013	-0.016	-0.024	-0.024
	(0.013)	(0.013)	(0.013)	(0.022)	(0.021)	(0.021)	(0.021)	(0.018)	(0.018)
Net Confidence Index$_{i,t}$	-0.008	-0.008	-0.008	-0.012	-0.013	-0.013	-0.012	-0.013	-0.013
	(0.011)	(0.011)	(0.011)	(0.015)	(0.015)	(0.015)	(0.021)	(0.021)	(0.021)
Exchange Rate$_{i,t}$	-0.041	-0.041	-0.041	0.052	0.049	0.049	-0.171	-0.171	-0.171
	(0.101)	(0.106)	(0.106)	(0.064)	(0.064)	(0.064)	(0.102)	(0.106)	(0.106)
Inflation (CPI)$_{i,t}$	1.124**	1.128**	1.128**	0.063	0.066	0.066	0.336	0.333	0.333
	(0.452)	(0.451)	(0.451)	(0.183)	(0.181)	(0.181)	(0.505)	(0.508)	(0.508)
Country Fixed Effects	x	x	x	x	x	x	x	x	x
Time Fixed Effects	x	x	x	x	x	x	x	x	x
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5318	5318	5318	4456	4456	4456	3040	3040	3040
R^2	0.168	0.173	0.173	0.105	0.108	0.108	0.145	0.148	0.148

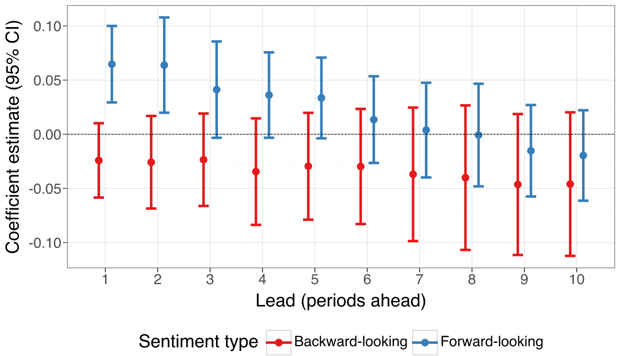
Notes: This table presents coefficient estimates from fixed-effects panel regressions where the dependent variable is the one-period-ahead change in the policy rate (Cols. I–III), the T-bill rate (Cols. IV–VI), and the T-bond rate (Cols. VII–IX). Net Policy Sentiment is based on the sentence-level classification of monetary policy decisions. Specs. (I), (IV), and (VII) show the total net policy sentiment (backward-looking + forward-looking); Specs. (II), (V), and (VIII) use the backward- and forward-looking components separately; and Specs. (III), (VI), and (IX) show the forward-looking and the gap between forward- and backward-looking components. Control variables include textual indicators and macroeconomic variables (CPI and exchange rate – USD/local). Dependent and independent variables are standardized. Country and time fixed effects are included. Standard errors clustered at the country level. Significance levels: * p <0.1, ** p <0.05, *** p <0.01.

Forward-looking net policy sentiment explains future changes in the policy rate in both advanced and other economies (Specs. I–II). This suggests that forward-looking statements serve as a reliable signal of central banks' policy intentions across various institutional and macroeconomic contexts.

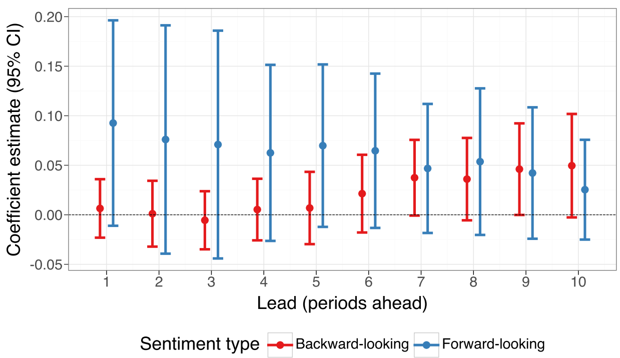
Figure 24: Relationship Between Forward-Looking Net Policy Sentiment and the Future Policy and Market Rates



(a) Changes in policy rates ahead



(b) Changes in T-bill interest rates (short-term) ahead



(c) Changes in T-bond interest rates (long-term) ahead

Notes: This figure presents estimated coefficients from panel regressions of future realized changes in policy rates and market interest rates on forward- and backward-looking net policy sentiment. Panel (a) shows results for policy rates, Panel (b) for short-term government securities (T-bills), and Panel (c) for long-term bonds (T-bonds). Each point corresponds to a separate regression at different leads, ranging from one to ten policy decisions ahead, following Equation (14). Vertical bars denote 95 percent confidence intervals.

For market-driven interest rates, however, the pattern of results diverges. Forward-looking sentiment significantly explains future movements in both short-term (T-bill) and long-term (T-bond) rates only in emerging and low-income economies. This heterogeneity suggests that forward-looking communication plays a more prominent role in shaping interest rate expectations in economies where financial markets are less developed, less liquid, or more information-constrained. In such contexts, central bank communication likely serves as a primary input for market participants, who depend more heavily on policy signals to form expectations. The ability of forward-looking sentiment to shift both short- and long-term yields underscores its central role in influencing expectations along the entire yield curve in these environments.

In advanced economies, by contrast, the weak or absent coefficients for forward-looking sentiment on market rates (Specs. IV and VI) are consistent with the hypothesis that financial markets in these jurisdictions more efficiently internalize and anticipate the behavior of central banks. The content of forward-looking communication may already be priced into market rates through a broader set of expectations mechanisms—such as macroeconomic forecasts, high-frequency financial data, and sophisticated modeling by private agents—which reduces the marginal information content of central bank statements. This interpretation aligns with the literature that emphasizes the "news content" of communication in emerging and less developed markets. In any case, the results are associations and should not be interpreted causally, as we lack an empirical identification strategy.

Table 8: Heterogeneous effects of the net policy sentiment on policy and interest rate changes by the economy's level of development

	Δ Policy Rate$_{i,t+1}$		Δ T-Bill Interest Rate$_{i,t+1}$		Δ T-Bond Interest Rate$_{i,t+1}$	
	Other Levels of Development	Advanced Economies	Other Levels of Development	Advanced Economies	Other Levels of Development	Advanced Economies
	(I)	(II)	(III)	(IV)	(V)	(VI)
Net Policy Sentiment (Forward)$_{i,t}$	0.079***	0.106**	0.058***	0.055	0.109*	-0.047
	(0.018)	(0.039)	(0.021)	(0.031)	(0.053)	(0.046)
Net Policy Sentiment (Backward)$_{i,t}$	0.064***	0.071**	-0.022	0.026	0.013	0.033
	(0.016)	(0.030)	(0.020)	(0.017)	(0.018)	(0.023)
Straightforwardness Index$_{i,t}$	-0.039*	0.020	-0.033**	-0.018	-0.027	0.020
	(0.020)	(0.031)	(0.016)	(0.020)	(0.023)	(0.014)
Explanation Index$_{i,t}$	0.000	0.017	-0.003	0.011	-0.016	-0.018
	(0.012)	(0.021)	(0.020)	(0.025)	(0.023)	(0.014)
Net Confidence Index$_{i,t}$	-0.009	-0.015	-0.011	0.002	-0.002	-0.040
	(0.015)	(0.016)	(0.017)	(0.023)	(0.029)	(0.025)
Exchange Rate (USD/domestic)$_{i,t}$	-0.033	-0.025	0.033	-0.141	-0.187*	-0.057
	(0.075)	(0.162)	(0.047)	(0.109)	(0.098)	(0.058)
Inflation (CPI)$_{i,t}$	1.158**	0.307	0.045	0.259	0.077	0.001
	(0.464)	(0.227)	(0.228)	(0.144)	(0.573)	(0.076)
Country Fixed Effects	x	x	x	x	x	x
Time Fixed Effects	x	x	x	x	x	x
Observations	3797	1518	3201	1248	1690	1349
R^2	0.188	0.422	0.158	0.334	0.210	0.516

Notes: This table presents coefficient estimates from regressions based on Equation (14), examining the effect of forward- and backward-looking net policy sentiment on future changes in policy rates, T-bill, and T-bond interest rates. The sample is divided into two groups based on the level of economic development: advanced economies and a combined group of emerging, developing, and low-income economies (denoted as "other levels of development"). Dependent variables are standardized changes in the monetary policy rate, short-term (T-bill), and long-term (T-bond) interest rates between t and $t+1$. The key explanatory variables are the forward- and backward-looking components of net policy sentiment. All regressions control for the straightforwardness, explanation, and net confidence indices, as well as inflation and exchange rate – USD/domestic. Country and time fixed effects are included. Standard errors clustered at the country level are shown in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Overnight index swap (OIS) rates and the net policy sentiment: We extend our analysis by examining how central bank communication shapes market expectations, measured through Overnight Indexed Swap (OIS) rates.[17] Unlike government bond yields or treasury bills, OIS rates are not subject to

[17] An Overnight Indexed Swap (OIS) is a derivative contract in which counterparties exchange a fixed interest rate

liquidity distortions or sovereign risk premia, making them a cleaner proxy for monetary policy expectations. Ideally, this would constitute our primary empirical specification. However, OIS data are available only for a limited set of countries with sufficiently developed derivatives markets. Consequently, the analysis is a robustness check to validate our baseline findings based on treasury instruments.

We estimate:

$$\text{OIS Rate}_{i,t+1} = \alpha_i + \lambda_t + \beta \text{ Net Policy Sentiment}_{i,t} + \lambda \text{ Policy Rate}_{i,t} + X'_{i,t}\gamma + \varepsilon_{i,t}, \tag{15}$$

where OIS Rate$_{i,t+1}$ is the OIS rate for economy i at time $t+1$, across four tenors: $1, 3, 6, 12$ months. The explanatory variables include forward-looking and backward-looking components of net policy sentiment, the lagged policy rate, and the same set of control variables as previous regressions. All variables are standardized to facilitate comparability. We add the usual country and time fixed effects and cluster errors at the country level.

Our dataset comprises 15 economies with sufficient liquidity and market depth to produce reliable OIS series sourced from Bloomberg: Australia, Canada, People's Republic of China, the Czech Republic, India, Indonesia, Japan, Malaysia, New Zealand, Republic of Poland, Russian Federation, Sweden, Switzerland, Thailand, and the United States. After merging communication indicators, macroeconomic variables, and policy rates, we retain 491 to 605 observations, depending on the tenor. The data span from January 2015 to February 2025. While more limited in coverage than our baseline sample, this subset includes a diverse mix of advanced and emerging market economies with floating exchange rates and active interest rate derivative markets.

Table 9 reports the results. Forward-looking net policy sentiment is positively associated with OIS rates at all horizons. The effect increases monotonically with tenor and peaks at the 12-month horizon, consistent with the notion that forward guidance influences expectations more strongly at longer maturities. These effects are identified conditional on current policy rates and macroeconomic fundamentals, implying that the tone of forward-looking communication shifts market expectations beyond what is already reflected in observable economic conditions.

By contrast, backward-looking sentiment is not statistically significant across tenors. This asymmetry highlights that markets do not react to retrospective assessments but rather respond to signals about future conditions. The policy rate explains OIS rates across the curve strongly and significantly, confirming its role as a key anchor for market expectations. Crucially, forward-looking sentiment remains significant even after conditioning on the current policy stance

for a floating rate tied to the daily overnight interbank rate, compounded over the contract term. The floating leg reflects the effective overnight rate—such as the federal funds rate or its equivalent in other jurisdictions—which is tightly controlled by the central bank through its operational framework. As a result, the OIS rate—the fixed rate that equates the expected value of payments on both legs—represents the market's expectation of the average policy rate over the swap's maturity. Unlike yields on government bonds or treasury bills, OIS rates are not influenced by credit risk, liquidity premia, or maturity-specific demand pressures. This makes them a clean, high-frequency measure of anticipated monetary policy, directly reflecting how markets interpret central bank signals regarding the future policy path.

and macroeconomic fundamentals, suggesting that communication conveys additional information relevant to the expected trajectory of policy.

Inflation is negatively associated with OIS rates at shorter maturities. This likely reflects market perceptions that periods of high inflation, typically coinciding with already elevated policy rates, may be followed by future easing. Supporting this interpretation, unreported regressions that interact inflation with the policy rate yield a negative and significant coefficient. At longer maturities, the influence of inflation dissipates, consistent with a shift in market focus toward forward guidance and broader macroeconomic dynamics.

Table 9: Effect of the Net Policy Sentiment on Future OIS Rates at Different Maturities

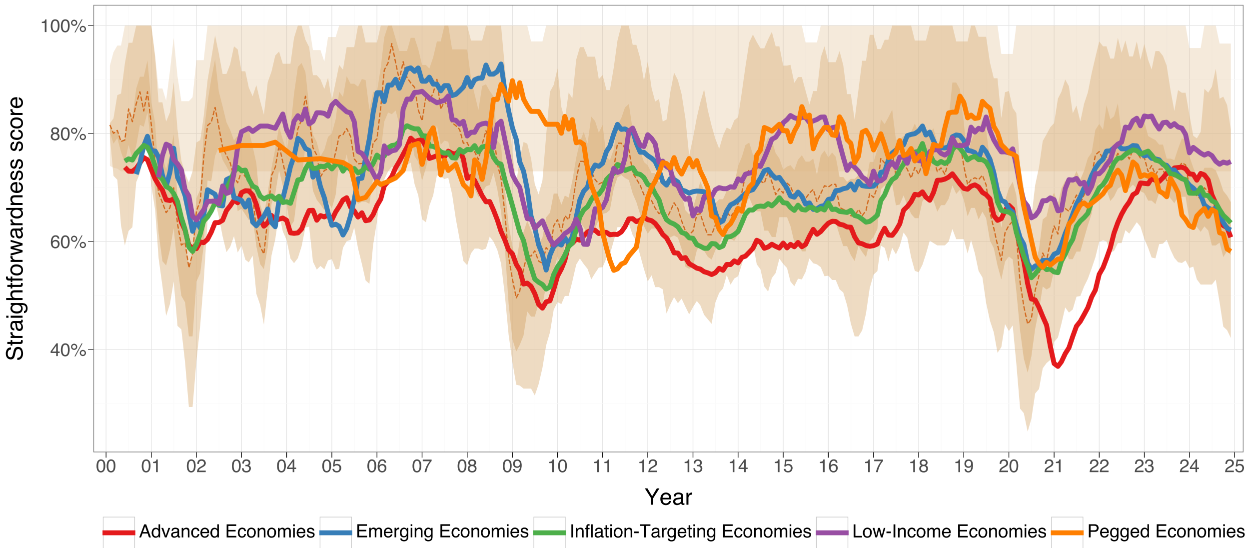
	OIS Rate$_{i,t+1}$ (1 month) (I)	OIS Rate$_{i,t+1}$ (3 months) (II)	OIS Rate$_{i,t+1}$ (6 months) (III)	OIS Rate$_{i,t+1}$ (12 months) (IV)
Net Policy Sentiment (Forward)$_{i,t}$	0.017**	0.024***	0.029***	0.036***
	(0.007)	(0.007)	(0.008)	(0.008)
Net Policy Sentiment (Backward)$_{i,t}$	0.017	0.032	0.034	0.036
	(0.013)	(0.020)	(0.025)	(0.028)
Straightforwardness Index$_{i,t}$	-0.014**	-0.012**	-0.008	-0.002
	(0.006)	(0.005)	(0.008)	(0.009)
Explanation Index$_{i,t}$	0.025*	0.014	0.004	-0.014
	(0.011)	(0.010)	(0.011)	(0.014)
Net Confidence Index$_{i,t}$	-0.001	-0.001	-0.007	-0.006
	(0.004)	(0.005)	(0.008)	(0.009)
Policy Rate$_{i,t}$	1.949***	1.951***	1.950***	1.838***
	(0.017)	(0.057)	(0.073)	(0.097)
Inflation (CPI)$_{i,t}$	-3.207***	-1.922***	-1.785**	-0.532
	(0.715)	(0.588)	(0.736)	(0.960)
Exchange Rate (USD/domestic)$_{i,t}$	-0.231	-0.164	-0.286	-0.165
	(0.193)	(0.199)	(0.259)	(0.192)
Country Fixed Effects	x	x	x	x
Time Fixed Effects	x	x	x	x
Observations	491	504	556	605
R^2	0.991	0.983	0.976	0.968

Notes: This table presents fixed-effects regression estimates of the net policy sentiment and control variables on the standardized OIS swap rate for different maturities (1m, 3m, 6m, 12m). The forward- and backward-looking net policy sentiment indices are derived from monetary policy communications. All variables are standardized. Controls include communication indices (straightforwardness, explanation, and net confidence indices), inflation (CPI), the policy rate, and the exchange rate. Country and time fixed effects are included. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.2.2. Straightforwardness Index

Figure 25 presents the evolution of the straightforwardness index in monetary policy decisions from 2000 to 2025. The shaded regions capture the interdecile range (10th to 90th percentile) across all economies in the sample, providing a view of cross-country dispersion over time. The dashed brown line indicates the global median. Colored lines represent the median values for groups of countries categorized by level of economic development and monetary policy regime.

Figure 25: Straightforwardness of Monetary Policy Communication



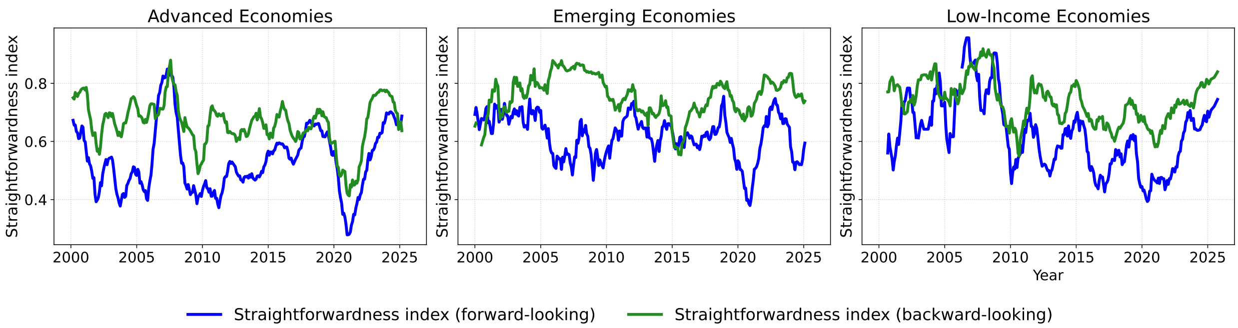
Notes: This figure shows the evolution of the average straightforwardness index in monetary policy decisions from 2000 to 2025 across economies with the same level of development and monetary policy framework. The shaded area represents the global interdecile range (10^{th}–90^{th} percentile), while the dashed brown line shows the global median. Colored lines correspond to median values by level of development (advanced, emerging, low-income) and monetary policy framework (inflation-targeting and pegged regimes).

Two key findings emerge. First, the drop in straightforwardness during episodes of systemic stress, such as the global financial crisis and the COVID-19 shock, is a pervasive and deliberate communication response across all groups of countries, regardless of their level of development or monetary policy framework. Central banks globally tend to reduce the unidirectionality of their messages during crises, shifting toward more nuanced and conditional communication. This pattern reflects a strategic effort to acknowledge uncertainty, articulate policy contingencies, and enhance transparency by expressing alternative scenarios and risk trade-offs. Rather than issuing unidirectional signals, central banks may use risk analyses to manage expectations and preserve credibility under volatile conditions. This finding aligns with the literature that studies the role of communication in conveying a robust policy reaction function under uncertainty (Blinder et al., 2008; Gertler & Karadi, 2015).

Second, the level of straightforwardness in monetary policy decisions outside of crises varies systematically across monetary and institutional settings. Advanced and inflation-targeting economies exhibit lower median straightforwardness, reflecting a communication strategy that prioritizes transparency of multiple potential risk scenarios in the future. In contrast, low-income and pegged exchange rate economies maintain higher straightforwardness scores in normal times, reflecting a stronger emphasis on clear, unambiguous messaging in the face of weaker nominal anchors, more significant exchange rate pass-through, and higher external vulnerabilities. From 2021 to 2024, straightforwardness scores increased steadily across most groups, likely reflecting a shift toward more unidirectional communication as central banks globally tightened policy rates in response to persistent inflation. The slight decline observed after 2024 suggests a return to more cautious messaging amid growing uncertainty about the global disinflation process and potential downside risks.

Figure 26 decomposes the straightforwardness index into forward- and backward-looking components across advanced, emerging, and low-income economies. Across all groups, backward-looking statements are consistently more straightforward than forward-looking ones, as expected given that past events are naturally described directly. Forward-looking communication, by contrast, tends to be less straightforward as it often conveys alternative paths and conditional scenarios. This empirical pattern reinforces the validity of the index: it captures the intuitive distinction between more unidirectional, factual communication about the past and the inherently more cautious and qualified nature of statements about the future.

Figure 26: Forward- and Backward-Looking Straightforwardness by Level of Development



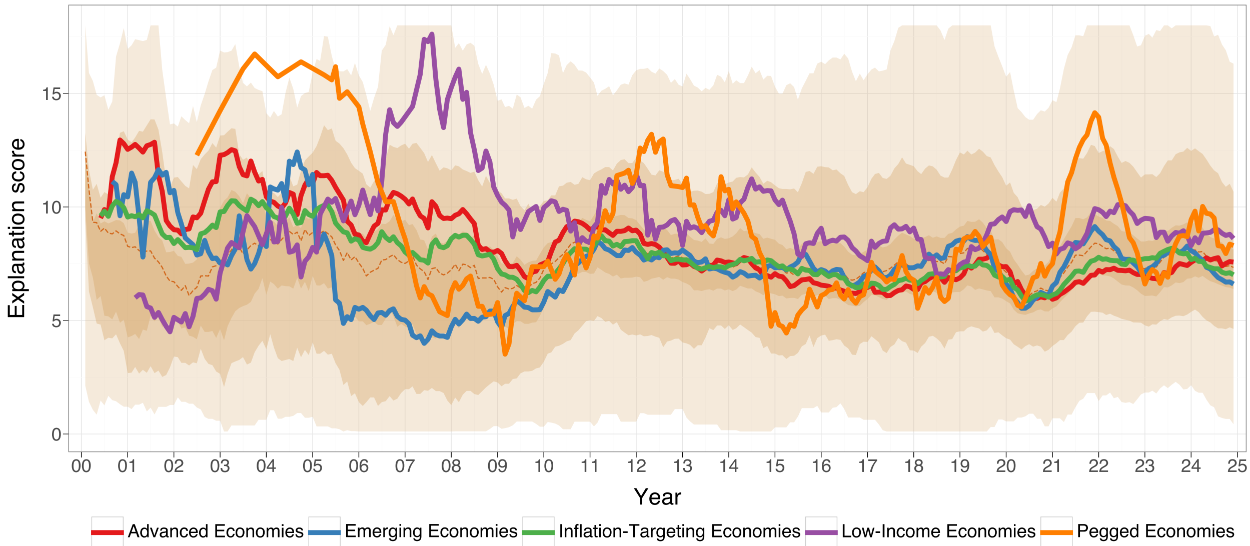
Notes: This figure shows the evolution of the straightforwardness index in monetary policy communication from 2000 to 2025, disaggregated into forward-looking (blue lines) and backward-looking (green lines) content. Each panel presents group-specific medians by level of development: advanced, emerging, and low-income economies.

5.2.3. Explanation Index

Figure 27 shows the evolution of the Explanation Index (EI) in monetary policy decisions from 2000 to 2025. Recall from Equation (9) that the EI is the sum of counts of confidence-building, risk-highlighting, and neutral statements normalized by the sum of hawkish and dovish statements. The shaded band represents the 10$^{\text{th}}$ to 90$^{\text{th}}$ percentile range across all countries, while the colored lines track the group medians across economies differentiated by level of development and monetary policy regime.

Explanation intensity rises systematically during periods of policy tightening, such as the post-COVID-19 period, when central banks implemented synchronized interest rate hikes. These phases are marked by intensified communication efforts aimed at justifying policy shifts and clarifying the underlying reaction function, consistent with the need to shape expectations and limit adverse market responses. In contrast, explanation scores tend to fall during easing cycles, notably at the onset of the COVID-19 crisis in 2020. This asymmetry reflects a well-documented pattern in central bank communication. Easing is generally perceived as a necessary and broadly accepted response to deteriorating conditions. However, tightening, especially after prolonged

Figure 27: Explanation in Monetary Policy Communication



Notes: This figure displays the evolution of the explanation index in monetary policy decisions from 2000 to 2025 across economies with the same level of development and monetary policy framework. The shaded area represents the global interdecile range (10^{th}–90^{th} percentile), and the dashed brown line shows the global median. Colored lines represent group medians by level of development (advanced, emerging, and low-income economies) and by monetary policy framework (inflation-targeting and pegged regimes).

accommodation, requires more explicit justification to avoid market overreaction (Campbell et al., 2012).

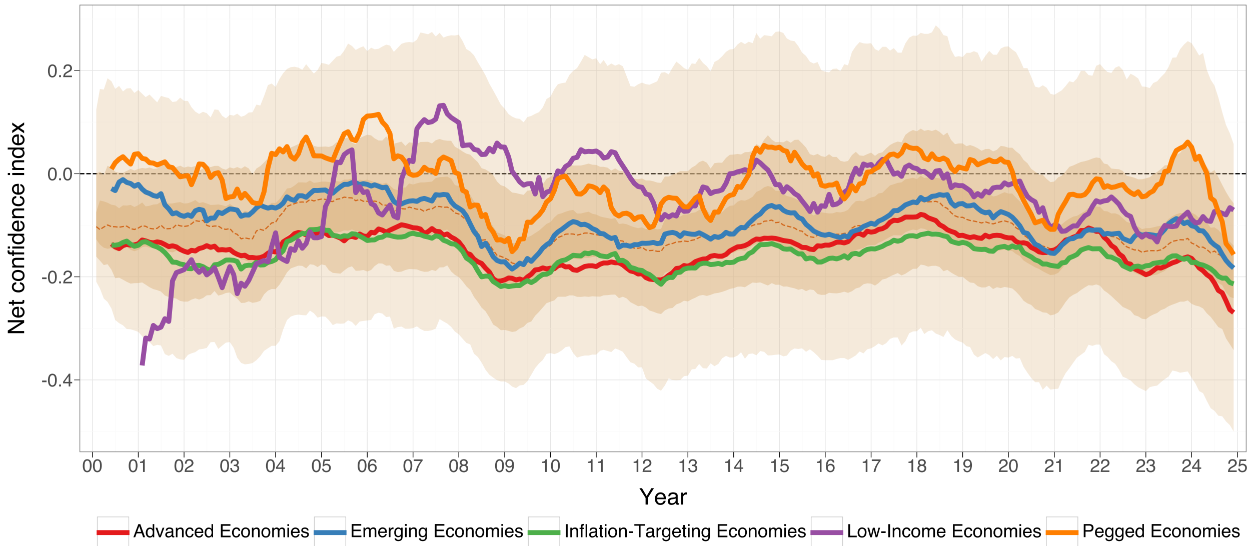
Although levels of the explanation index vary across country groups, its movements are synchronized among most groups, underscoring the dominant role of global shocks in shaping communication dynamics. Low-income and pegged exchange rate economies tend to exhibit higher explanation scores, likely reflecting the need to compensate for weaker institutional credibility or limited monetary policy flexibility. However, cross-group differences have narrowed in recent years. Since 2015, the index has remained within a relatively tight range, with transitory spikes coinciding with major turning points in the global policy cycle. These patterns suggest that explanation intensity is not a structural attribute of institutional development but rather a responsive feature of communication strategy, deployed more forcefully when policy tightening introduces uncertainty or reverses established guidance.

5.2.4. Net Confidence Index

Figure 28 depicts the evolution of the net confidence index from 2000 to 2025, based on the full suite of regular central bank publications, including monetary policy decisions and reports, financial stability reports, annual reports, and speeches. Unlike the previous measures that focused specifically on monetary policy decisions, this index captures risk communication strategies, reflecting how central banks balance expressions of confidence with the highlighting of vulnerabilities in their communications.

The index is persistently negative across most country groups and periods, indicating that central

Figure 28: Net Confidence Index (Risk Communication) in Central Bank Communication



Notes: This figure shows the evolution of the net confidence index (risk communication) in central bank communications from 2000 to 2025 across economies with the same level of development and monetary policy framework. The index is constructed from sentence-level classifications of all major regular publications, including monetary policy reports, financial stability reports, annual reports, and speeches. The shaded area represents the global interdecile range (10^{th}–90^{th} percentile). Colored lines represent median values by level of development (advanced, emerging, and low-income economies) and by monetary policy framework (inflation-targeting and pegged regimes).

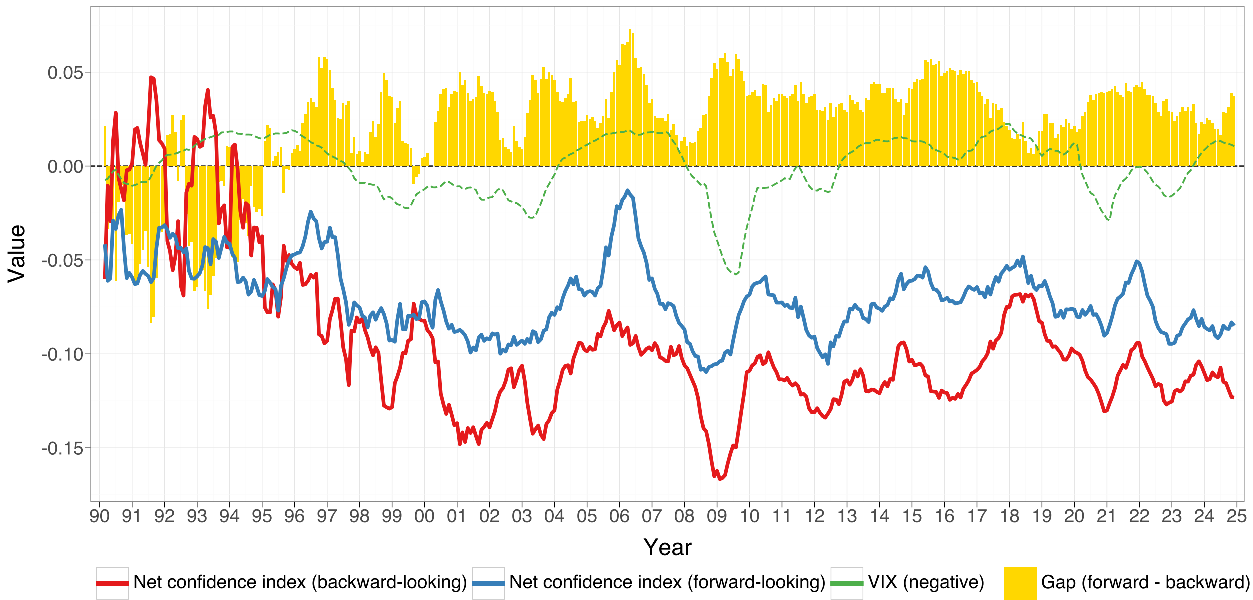
banks prioritize highlighting risks over building confidence in their regular communications. This behavior likely reflects an institutional strategy of risk awareness: by highlighting vulnerabilities, central banks reinforce their role as overseers of financial stability and signal vigilance in the face of uncertainty. This approach is particularly pronounced in advanced and inflation-targeting economies. Pegged economies exhibit the highest volatility in the net confidence index. This reflects their heightened exposure to external shocks and the frequent need to recalibrate communication in response to exchange rate pressures, capital flow dynamics, or geopolitical developments—factors that inherently move exchange rates, their nominal anchor.

Importantly, the net confidence index also tracks the underlying risk environment in the global economy. It declines sharply during episodes of systemic stress, such as the global financial crisis and the COVID-19 pandemic. Conversely, periods of economic recovery tend to be associated with more neutral or mildly positive net confidence values. Thus, the index serves as a barometer of perceived macro-financial risk, reflecting not only the tone of central bank communication but also its responsiveness to changing economic conditions.

While the aggregate net confidence index offers valuable insights into how central banks balance confidence-building and risk communication, further analytical depth can be gained by examining two additional dimensions. The first relates to the temporal framing of risks: distinguishing between forward- and backward-looking statements is essential to understanding how central banks shape expectations about future conditions versus reflecting on prevailing circumstances. The second concerns the global relevance of communication patterns. Aggregating national indices to a global level allows us to assess how central banks collectively frame risks and confidence, with particular attention to the systemic role played by large and interconnected economies.

Figure 29 presents the forward- and backward-looking components of the net confidence index aggregated at the global level from 1990 to 2025. Aggregation is performed using GDP weights in US dollars to reflect the relative influence of each economy. The red line represents the backward-looking net confidence index, while the blue line shows the forward-looking component. The yellow bars illustrate the gap between the two; positive values indicate that central banks express more confidence in future conditions than in current or past ones, providing a measure of directional risk communication. The dashed green line represents the negative of the VIX index to facilitate comparison, as a higher VIX signals greater uncertainty. The series begins in 1990, the earliest year with VIX data available.

Figure 29: Forward- and Backward-Looking Net Confidence vs. Implied Market Volatility



Notes: This figure compares the global GDP-weighted net confidence index—split into forward-looking (blue) and backward-looking (red) components—with the inverse of the VIX index (green dashed line) from 1990 to 2025. The backward-looking index reflects assessments of current and recent macro-financial conditions, while the forward-looking index captures expectations about future stability. Yellow bars represent the gap between forward- and backward-looking sentiment, with positive values indicating stronger forward confidence. The VIX is standardized and inverted to align visually with the direction of the net confidence index (i.e., higher risk perception corresponds to lower confidence).

An interesting feature is that central banks tend to express more confidence in their forward-looking communication than their backward-looking assessments, at least over the past 30 years of communication. This does not contradict earlier findings that central banks generally emphasize risks; rather, it highlights a relative shift in tone: even when they highlight current vulnerabilities, they often frame future conditions with cautious relative optimism considering current conditions. This optimistic forward tilt may reflect strategic communication objectives. By presenting a more confident outlook, central banks aim to stabilize expectations, reinforce perceptions of policy efficacy, and minimize the risk of amplifying uncertainty—an approach aligned with theoretical models that emphasize the critical role of communication in shaping private sector behavior under incomplete information (Woodford, 2005).

Connection between the VIX and the net confidence index: The observed co-movement between the net confidence index and the VIX highlights a potential two-way interaction between central bank communication and financial market volatility. Central banks use communication as a tool to influence private sector expectations, with substantial empirical evidence showing that their statements can move financial markets even beyond the effects of realized policy actions (e.g., Sturm & De Haan; Woodford (2011; 2005)). Communication about financial stability issues, in particular, has been found to affect asset prices, especially during episodes of heightened stress. In this line, optimistic financial stability reports can lead to positive stock market reactions (e.g., Born et al. (2014)). Furthermore, empirical evidence suggests that substantial changes in the content of central bank statements, following periods of highly similar communications, can increase market volatility, indicating that markets are sensitive not only to policy decisions but also to the way central banks frame risks and uncertainties (e.g., Ehrmann & Talmi (2020)).

While these findings establish that communication significantly affects market behavior, they leave open the question of directionality: do central banks adapt their communication in response to shifts in market volatility, or do forward-looking statements actively shape future risk perceptions? Addressing this question requires an econometric approach that can disentangle the sequencing of information flows. To this end, we examine the connection between the Net Confidence Index and the VIX using Granger causality analysis. Unlike the previous analysis of Net Policy Sentiment, which focused on conditional correlations using fixed-effects panel regressions, this research interest specifically lies in the timing and direction of predictive relationships between two evolving global aggregates. Although lead-lag panel regressions could offer some insight into temporal associations, they do not formally test whether past values of one variable improve the prediction of another beyond the information contained in the past values of the dependent variable itself. Granger causality analysis addresses this limitation by providing a structured statistical test of incremental predictive content, assessing whether forward-looking communication contains signals about future financial conditions or whether market volatility precedes changes in backward-looking communication. While Granger causality does not recover structural causal effects and does not control for potential confounders, it is well-suited for establishing the temporal ordering of informational flows, which is the focus of this part of the analysis.

The Granger causality tests use three monthly global time series aggregated from 1990 to 2025: the forward-looking and backward-looking components of the Net Confidence Index, and the standardized VIX. Each Net Confidence Index component is computed as a GDP-weighted average across countries, reflecting differences in economic size. The resulting dataset contains 420 monthly observations for each series. Summary statistics are presented in Table 2 in Appendix E.

We implement bivariate Granger causality tests separately for each pairwise relationship: (i) forward-looking confidence and the VIX, and (ii) backward-looking confidence and the VIX. The maximum number of lags considered is twelve months, allowing for the possibility that adjustments in communication or financial market volatility may occur with some delay. Granger causality formally tests whether the inclusion of lagged values of a predictor improves the out-of-sample prediction of a target variable, relative to a model that relies only on the target's own lags.[18] Although Granger causality captures predictive content and temporal ordering, it does not establish causality in the economic sense, as it does not control for potential omitted variables or contemporaneous confounders. All the global series are treated to ensure stationarity before the Granger causality analysis.[19] Table 10 reports the Granger causality results.

Panel A shows that past values of the forward-looking Net Confidence Index significantly improve the prediction of future realized VIX, with statistically significant F-statistics up to six lags. Rather than implying causality, these results reveal predictive content: forward-looking central bank communication carries signals about upcoming fluctuations in market volatility. This finding supports the view that proactive risk communication can play a meaningful role in shaping future financial market risk perceptions, consistent with the notion that central banks influence not only macroeconomic expectations but also financial risk sentiment. Panel B, by contrast, finds no evidence that lagged VIX values predict changes in forward-looking confidence. This is consistent with the idea that the forward-looking Net Confidence Index reflects policymakers' assessments of current and prospective risk conditions, rather than mechanically reacting to past market volatility. Since the VIX primarily captures contemporaneous or expected near-term financial stress, its lagged values provide limited additional information for shaping forward-looking risk communication.

Panels C and D present the results connecting VIX and the backward-looking Net Confidence Index. The empirical relationship is reversed: Panel C shows that backward-looking communication has no predictive power for future VIX movements. At the same time, Panel D reveals that past VIX Granger-causes subsequent adjustments in backward-looking confidence, particularly two to seven months after a volatility shock. These results provide empirical evidence that the backward-looking Net Confidence Index captures ex-post assessments of realized risk environments. Moreover, the lag structure—where backward-looking narratives react with a delay to market stress—suggests that retrospective communication integrates financial market information gradually.

Taken together, these findings deepen our understanding of the informational content of central bank risk communication. Prior studies (e.g., Born et al.; Ehrmann & Talmi (2014; 2020))

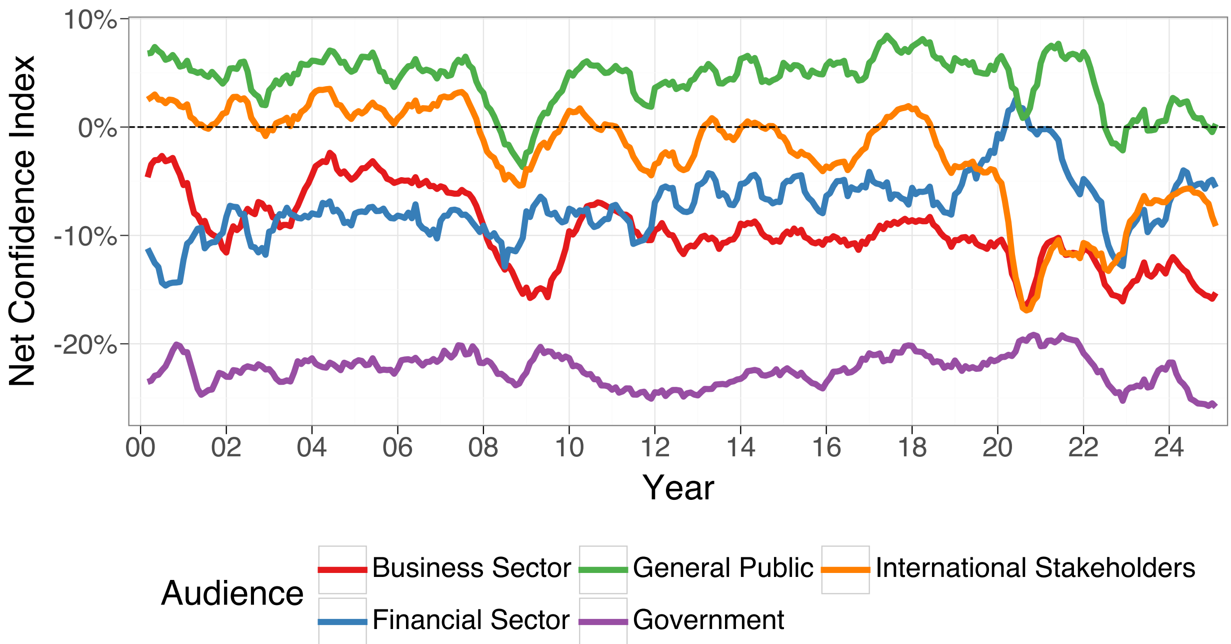
[18] Formally, for two stationary series X_t and Y_t, the Granger causality test estimates a system of equations of the form: $Y_t = \alpha + \sum_{k=1}^{p} \beta_k Y_{t-k} + \sum_{k=1}^{p} \gamma_k X_{t-k} + \epsilon_t$, where p is the number of lags. The null hypothesis is that all coefficients $\gamma_k = 0$ for $k = 1, \dots, p$, meaning that lagged values of X do not help predict Y beyond lagged values of Y itself. The test statistic is based on an F-test comparing the restricted model (which excludes lagged X) with the unrestricted model (which includes lagged X). Rejection of the null indicates that X Granger-causes Y.

[19] Before proceeding to Granger causality testing, we assess the stationarity properties of the time series to avoid spurious inferences. We apply both the Augmented Dickey-Fuller (ADF) test, which tests the null hypothesis of a unit root (nonstationarity), and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, which tests the null hypothesis of stationarity. For the forward-looking Net Confidence Index, the ADF test fails to reject nonstationarity ($p = 0.265$), and the KPSS test rejects stationarity ($p = 0.010$), providing consistent evidence of nonstationarity. For the backward-looking component, the ADF test rejects the unit root null ($p = 0.013$), but the KPSS test simultaneously rejects stationarity ($p = 0.010$), yielding mixed evidence. For the VIX, both tests point toward stationarity (ADF $p = 0.007$; KPSS $p = 0.290$). Given these findings, we use the forward- and backward-looking Net Confidence Index series in differences to ensure stationarity, while the VIX is kept in levels. This approach ensures that Granger causality inferences are drawn from approximately stationary series.

have established that financial stability communication can influence market behavior, particularly during episodes of distress. However, our results advance the literature by documenting that distinguishing between proactive and reactive communication components at the sentence level not only enhances the precision of narrative measurement but also reveals distinct transmission patterns related to financial volatility. This highlights the importance of disaggregating communication strategies when assessing their impact on financial market dynamics.

Central banks tailor sentiment to the audience. We compute the net confidence index separately for each audience to assess how central banks calibrate their communication across audiences. Specifically, for every document, we calculate five distinct net confidence indices—one for each of the key audiences: general public, government, business sector, financial sector, and international stakeholders. Figure 30 presents the evolution of these audience-specific indices from 2000 to 2025. We start our analysis in the 2000s because it marks the decade when central banks broadened their communication practices beyond annual reports to include regular monetary policy statements, financial stability publications, and various press materials, thus enabling a more granular analysis of communication strategies.

Figure 30: Tailoring Sentiment to Different Audiences in Central Bank Communication



Notes: This figure shows the evolution of the net confidence index in central bank communication from 2000 to 2025, disaggregated by audience (main recipient of the message): general public, government, business sector, financial sector, and international stakeholders. For each document, the index is computed by applying the net confidence formula only to the subset of sentences classified as addressing a given audience. Positive values indicate a predominance of confidence-building language; negative values reflect a greater emphasis on risk.

Our empirical results reveal no uniform or generic communication style across audiences. Instead, central banks consistently tailor their messages to reflect the informational needs, expertise,

Table 10: Granger Causality Tests: Differenced Net Confidence Indices (NCI) and VIX

Panel A: Δ Forward-looking NCI → VIX

Lag	F-stat	p-value	Signif.	Direction
1	0.417	0.519		Δ Forward → VIX
2	3.563	0.029	**	Δ Forward → VIX
3	1.840	0.139		Δ Forward → VIX
4	3.151	0.014	**	Δ Forward → VIX
5	1.980	0.081	*	Δ Forward → VIX
6	2.013	0.063	*	Δ Forward → VIX
7	1.473	0.175		Δ Forward → VIX
8	1.313	0.235		Δ Forward → VIX
9	1.205	0.290		Δ Forward → VIX
10	1.205	0.286		Δ Forward → VIX
11	1.217	0.273		Δ Forward → VIX
12	1.161	0.309		Δ Forward → VIX

Panel B: VIX → Δ Forward-looking NCI

Lag	F-stat	p-value	Signif.	Direction
1	0.719	0.397		VIX → Δ Forward
2	2.074	0.127		VIX → Δ Forward
3	1.973	0.117		VIX → Δ Forward
4	1.387	0.237		VIX → Δ Forward
5	0.964	0.440		VIX → Δ Forward
6	0.833	0.545		VIX → Δ Forward
7	0.803	0.585		VIX → Δ Forward
8	0.682	0.708		VIX → Δ Forward
9	0.853	0.567		VIX → Δ Forward
10	1.000	0.443		VIX → Δ Forward
11	0.976	0.467		VIX → Δ Forward
12	0.921	0.525		VIX → Δ Forward

Panel C: Δ Backward-looking NCI → VIX

Lag	F-stat	p-value	Signif.	Direction
1	6.230	0.013	**	Δ Backward → VIX
2	0.449	0.638		Δ Backward → VIX
3	0.598	0.616		Δ Backward → VIX
4	0.498	0.737		Δ Backward → VIX
5	0.373	0.867		Δ Backward → VIX
6	0.224	0.969		Δ Backward → VIX
7	0.869	0.531		Δ Backward → VIX
8	0.793	0.609		Δ Backward → VIX
9	0.750	0.663		Δ Backward → VIX
10	0.780	0.648		Δ Backward → VIX
11	0.843	0.597		Δ Backward → VIX
12	0.687	0.764		Δ Backward → VIX

Panel D: VIX → Δ Backward-looking NCI

Lag	F-stat	p-value	Signif.	Direction
1	0.762	0.383		VIX → Δ Backward
2	3.366	0.035	**	VIX → Δ Backward
3	3.293	0.021	**	VIX → Δ Backward
4	2.717	0.029	**	VIX → Δ Backward
5	2.279	0.046	**	VIX → Δ Backward
6	2.255	0.037	**	VIX → Δ Backward
7	1.867	0.073	*	VIX → Δ Backward
8	1.394	0.197		VIX → Δ Backward
9	1.242	0.268		VIX → Δ Backward
10	1.347	0.203		VIX → Δ Backward
11	1.354	0.193		VIX → Δ Backward
12	1.180	0.295		VIX → Δ Backward

Notes: This table reports bivariate Granger causality tests examining whether lagged values of the differenced Net Confidence Index components (forward- and backward-looking) improve the prediction of the VIX, and vice versa. Panels A and B present results for the forward-looking component; Panels C and D present results for the backward-looking component. Each panel reports F-statistics and p-values for tests of the joint significance of twelve monthly lags of the predictor. The differenced series of the forward- and backward-looking components of net confidence indices are used to ensure stationarity. *** p<0.01, ** p<0.05, * p<0.1.

and policy relevance of each group. Communication directed at the general public is systematically framed in the most confident and reassuring terms. This likely reflects the objective of promoting trust and reducing uncertainty among a broad audience that tends to be less financially sophisticated and more vulnerable to shifts in macroeconomic conditions and sentiment. By conveying a sense of confidence and stability, central banks help anchor public expectations and support economic behavior aligned with policy objectives.

By contrast, communication targeting the government consistently emphasizes risks more heavily, reflected in its markedly lower confidence scores. This risk-oriented tone may reflect the central bank's need to signal fiscal prudence and highlight macroeconomic vulnerabilities, which are directly relevant to monetary policy and financial stability issues. Messages to the business and financial sectors, while also on the risk-aware side, exhibit a comparatively more neutral and balanced tone. These audiences respond more actively to economic signals when making investment and financing decisions. However, their sophisticated analytical capacities may incentivize central banks to favor clarity, precision, and fact-based communication over excessively optimistic or pessimistic language. In these cases, neutral communication helps avoid misinterpretation or market overreaction.

The empirical analysis also shows a distinct pattern of central bank communication during distressed times compared to normal times. During periods of systemic stress, such as the global financial crisis, the COVID-19 shock, and episodes of market turbulence, central banks communicate by building relatively more confidence in the financial sector and the government. This likely reflects a deliberate effort to reassure key institutional actors in crisis mitigation. The financial sector, in particular, serves as the primary transmission mechanism for central bank policy actions. Reinforcing confidence within this sector during crises is essential to ensure continued credit provision, liquidity management, and market functioning. At the same time, confidence in messages directed at the general public, business sector, and international stakeholders tends to decline during crises. This shift suggests a more cautious and transparent approach to these segments, where central banks emphasize risk awareness.

Perhaps most notably, these audience-specific patterns are not merely cyclical. Rather, the analysis indicates that the tailoring of sentiment across audiences is structural and persistent. Despite profound macroeconomic shocks and shifts in monetary regimes over the past three decades, the relative tone used for most audiences has remained remarkably stable.[20] This consistency points to intentional communication design, where differentiation in tone is embedded in central bank strategy rather than emerging as an ad hoc response to prevailing economic conditions.

We evaluate this hypothesis empirically using a formal variance decomposition exercise. Specifically, we assess how much of the variation in audience-specific sentiment is attributable to structural factors—such as systematic differences across audiences—versus transitory or contextual

[20] Communication addressed to international stakeholders exhibits somewhat more variation over time than communication addressed to other audiences. Significant variations occur in distressed periods, when sentiment directed at this audience becomes more negative, likely reflecting the increased global uncertainty. Thus, while structural differences dominate, international communication appears slightly more sensitive to global shocks, possibly due to its inherently outward-facing and interconnected nature.

components, such as country—specific and time-specific influences. To do so, we estimate a series of panel regressions with progressively richer sets of dummy variables. For each document and audience, we regress the audience-specific net confidence index on dummies capturing audience, country, communication outlet, and time dimensions. The contribution of each factor to overall variation is then quantified based on its share of explained variance.[21]

If central banks adopt persistent and systematic differences in tone across audiences, we would expect audience dummies to account for a large share of the explained variation. Conversely, if tone is heavily shaped by country-specific circumstances or cyclical developments, the corresponding country and time fixed effects would capture more of the variation. Finally, if much of the variation were idiosyncratic or noisy, the residual component would dominate.

Figure 31 shows the explained share for each dimension for three panel specifications with different data aggregations. We observe that the audience differentiation is an essential feature of central bank communication. When considering only time and audience fixed effects, differences across audiences explain the largest share of the variation in the net confidence index. Adding country and communication outlet dummies naturally reduces the portion attributable to audience fixed effects, but this contribution remains substantial across all model specifications. In contrast, time fixed effects consistently explain only a modest fraction of the variation, suggesting that cyclical factors and global shocks play a secondary role relative to the systematic differentiation in communication strategies. Similarly, the relatively small explanatory power of country fixed effects suggests that audience tone is not strongly shaped by idiosyncratic national conditions.

Taken together, these findings reinforce the interpretation that central banks' tone toward distinct audiences is shaped primarily by structural and intentional design choices, rather than by transitory macroeconomic or country-specific factors. Tone differentiation thus emerges as a systematic and persistent feature of modern central bank communication.

6. Conclusions

This research develops a novel automated classification framework to systematically analyze central bank communication, addressing the critical need for measurable and interpretable metrics in monetary policy discourse. Our framework operates at the sentence level, classifying text along four key dimensions—topic, communication stance, sentiment, and audience, offering

[21] Formally, the following model is estimated for each specification:

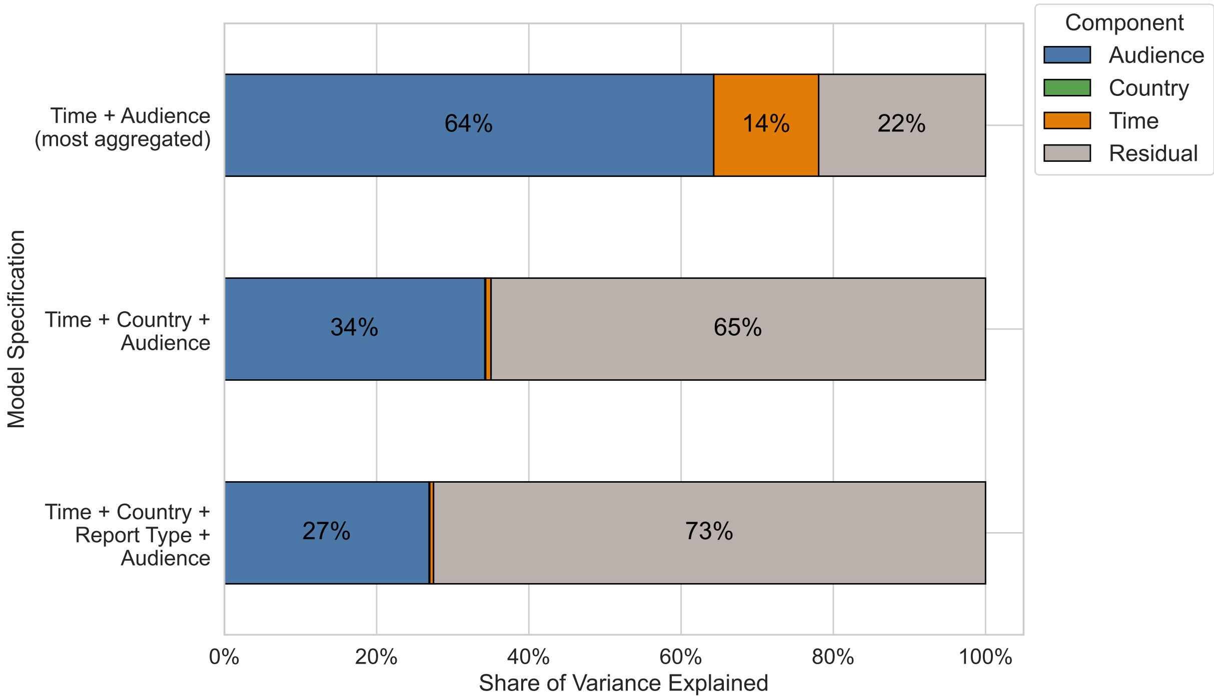
$$\text{Net Confidence Index}_{i,a,c,t} = \alpha + \sum_{w_i} \delta_{w_i} C^{w_i} + \sum_{w_a} \gamma_{w_a} A^{w_a} + \sum_{w_c} \rho_{w_c} R^{w_c} + \sum_{w_t} \theta_{w_t} T^{w_t} + \epsilon_{i,a,c,t},$$

in which i, a, c, and t index country, audience, communication outlet, and time, respectively. The terms A are audience dummies, C are country dummies, R are communication outlet dummies, and T are time dummies indexed over their full estimable range. The share of variance explained by each component is computed as:

$$\text{Share}_k = \frac{\text{SS}_k}{\text{SS}_{\text{Total}}},$$

where SS_k denotes the sum of squares attributable to component k, obtained from the model decomposition of the regression residuals, and SS_{Total} is the total sum of squares.

Figure 31: Audience Drives Most of the Variation in Central Bank Confidence Communication



Notes: This figure reports a variance decomposition of the net confidence index, computed separately for each audience. Bars show the share of variance explained by audience, country, and time dimensions, based on sequential fixed-effects panel regressions using progressively granular specifications. The audience dimension consistently accounts for the largest share of explained variance, underscoring the strategic tailoring of tone across stakeholder groups. Time effects explain little, suggesting global shocks and cyclical conditions are secondary. Country effects also contribute modestly, indicating limited influence from national-specific factors. The residual captures unexplained heterogeneity.

an unprecedented level of granularity in assessing policy messaging. While all dimensions are essential for understanding central bank communication, the explicit measurement of communication stance using textual information has been largely overlooked despite its central role in shaping expectations and influencing market behavior.

By leveraging a fine-tuned multilingual large language model trained explicitly for central bank communication, our approach captures semantic and contextual nuances that traditional dictionary-based methods fail to detect. By extracting policy signals with high precision, this framework advances the study of central bank transparency, allowing for a more rigorous evaluation of communication effectiveness.

Applying this framework to an extensive dataset of 74,882 documents from 169 central banks worldwide, we provide novel insights into the evolution of central bank messaging across time, countries, and monetary policy regimes. This is a treasure trove of data, and given the volume and extent we have only started to explore the insights that can be gleaned from this. Our findings reveal significant shifts in monetary policy communication when countries adopt inflation-targeting frameworks, with backward-looking exchange rate statements giving way to forward-looking discussions on inflation, interest rates, and economic activity. The framework also highlights how central banks tailor their messaging across different audiences, including financial markets, businesses, households, and international stakeholders, reflecting strategic adjustments in communication.

We introduce a suite of novel textual metrics with clear economic interpretations to complement the framework. The net policy sentiment metric quantifies the overall stance of communication, distinguishing between forward- and backward-looking components, with the former serving as a proxy for forward guidance. Our empirical analysis demonstrates that forward-looking sentiment robustly predicts policy rate changes and influences market interest rates, reinforcing its role as a monetary policy tool. Additionally, we introduce the straightforwardness index and the explanation index, which evaluate the clarity and depth of policy justifications. Furthermore, the net confidence index captures the balance between confidence-building and risk-highlighting statements, offering insights into how central banks manage uncertainty communication.

Beyond its academic contributions, this framework offers central banks practical tools to enhance accountability of past actions, build transparency of current actions, and shape expectations. Enabling policymakers to assess and refine their messaging systematically provides a means to benchmark communication against historical trends or peer institutions and align strategies with best practices. This feature is valuable for strengthening market confidence and improving monetary policy transmission. More broadly, our approach represents a significant step in translating qualitative text into quantitative data, enabling central banks and researchers to more systematically assess how communication shapes expectations, influences financial conditions, and interacts with policy decisions.

Future research could empirically examine the impact of communication on macroeconomic variables, particularly by analyzing our proxy for forward guidance in monetary policy decisions. One compelling direction is to investigate whether communication can influence expectations in a way that enhances monetary policy transmission and, consequently, contributes to price stability, a key concern for policymakers. In other words, this research could explore the potential of communication as a causal factor in inflation. Beyond monetary policy, the framework opens up new avenues to assess communication effectiveness in other critical areas of central bank mandates. For example, the net confidence index could be related to financial stress indicators, such as country-level financial stress indices, to understand better how risk communication shapes market sentiment and stability perceptions. Likewise, FX-related communication topics could be studied alongside exchange rate dynamics to investigate whether central bank messaging affects currency movements and volatility. Another promising extension is the application of this methodology to social media and other non-traditional communication channels, where timely and targeted messaging increasingly plays a role in shaping expectations. Addressing these areas will deepen our understanding of the complex dynamics between central bank communication and economic outcomes, while expanding the framework's applicability beyond monetary policy to broader aspects of financial stability and institutional accountability.

Appendix A. Prompts

This appendix reports the auxiliary prompts used throughout the paper. All prompts were run in the OpenAI API interface using the ChatGPT-4o chatbot model (as of April 2025).

A.1 Prompt: Generate Synthetic Examples

You are assisting in creating a supervised training dataset for classifying sentences from central bank communications in downstream tasks. Each sentence must be labeled across four dimensions:

- **Topic**: [Insert full list of topic labels here]
- **Communication Stance**: [forward-looking, backward-looking]
- **Audience**: [financial sector, business sector, households, government, international stakeholders]
- **Sentiment**: [hawkish, dovish, neutral/balanced, risk-highlighting, confidence-building]

You must adhere to two essential design principles:

1. **Semantic Differentiation Across Labels**: Ensure that sentences from different classes—particularly across topics and sentiment—are semantically distinct in subject matter, policy framing, and economic context. Avoid overlap between categories.
2. **Intra-Label Diversity**: For each label combination, vary the linguistic style, syntactic structure, and policy context while keeping the sentence firmly representative of its label set. This includes using different tones (e.g., formal, direct), constructions (e.g., passive/active), and framing devices.

Task: Given the label set provided below, generate `[X]` example sentences. Each sentence should:

- Reflect the tone and structure of authentic central bank communication
- Be labeled across the four classification dimensions
- Satisfy both semantic differentiation and intra-label diversity

Output format: Return only the labels, without explanation, in a JSON-compliant format where keys represent the dimension name and values represent the predicted label.

Label combination to generate:

- Topic: `[Your chosen topic label]`
- Communication Stance: `[Your chosen stance]`
- Audience: `[Your chosen audience]`
- Sentiment: `[Your chosen sentiment]`

Now, generate `[X]` sentences matching the above label combination.

A.2 Prompt: ChatGPT 4o Weakly-Supervised Classification

You are an assistant trained to analyze central bank communication. Please classify the following sentence along four dimensions:

- **Topic**: [Insert full list of topic labels here]
- **Communication Stance**: [forward-looking, backward-looking]
- **Audience**: [financial sector, business sector, households, government, international stakeholders]
- **Sentiment**: [hawkish, dovish, neutral/balanced, risk-highlighting, confidence-building]

Output format: Return only the labels, without explanation, in a JSON-compliant format where keys represent the dimension name and values represent the predicted label.

Task: Sentence to be classified: `[Sentence]`

Appendix B. Label Consistency Across Multiple Languages

This appendix assesses the robustness of our classification framework to multilingual inputs by measuring the consistency of predicted label distributions across original central bank communications and their corresponding translations. To conduct this evaluation, we selected a representative sample of non-English central bank documents across seven languages, covering the full spectrum of publication types—annual reports, financial stability reports, monetary policy reports, and monetary policy decisions. The sample includes 74 documents in Portuguese (PT), 37 in French (FR), 69 in Russian (RU), 19 in Romanian (RO), 16 in Spanish (ES), and 7 in Arabic (AR). Each document was translated into English using ChatGPT 4o. We then applied our fine-tuned sentence-level classifiers independently to both the original and translated versions, generating predicted label distributions for each document across four dimensions: topic, communication stance, audience, and sentiment.

To quantify consistency, we compared the distributions of predicted labels for each original-translation document pair using the Jensen–Shannon Divergence – JSD (Lin, 1991). The JSD is a symmetric measure of distributional difference derived from the Kullback–Leibler (KL) divergence. Formally, let $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ denote two discrete probability distributions (with n distinct labels). The JSD is defined as:

$$\mathrm{JSD}(P \parallel Q) = \frac{1}{2} D_{KL}(P \parallel M) + \frac{1}{2} D_{KL}(Q \parallel M),$$

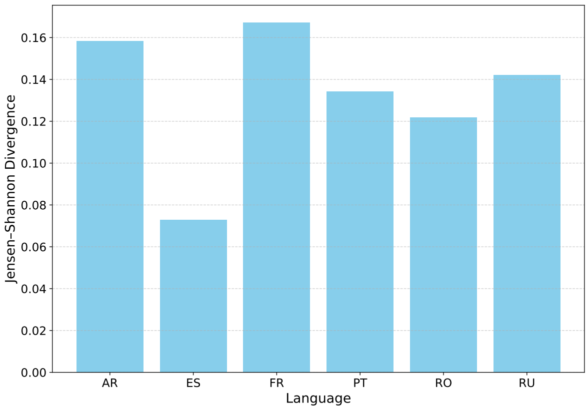
where $M = (P+Q)/2$ is the midpoint distribution, and $D_{KL}(P \parallel Q) = \sum_{i=1}^{n} p_i \log_2 \left(\frac{p_i}{q_i}\right)$ denotes the KL divergence from distribution P to Q. Because it is bounded between 0 (identical distributions) and 1 (maximally different), the JSD provides a meaningful and interpretable metric of divergence.

Figure 1 displays the average JSD across language pairs for each classification dimension. The observed divergences are remarkably low, with values rarely exceeding 0.15. This suggests a high degree of semantic consistency between the original and translated versions of central bank communications, underscoring the robustness of our multilingual classification framework.

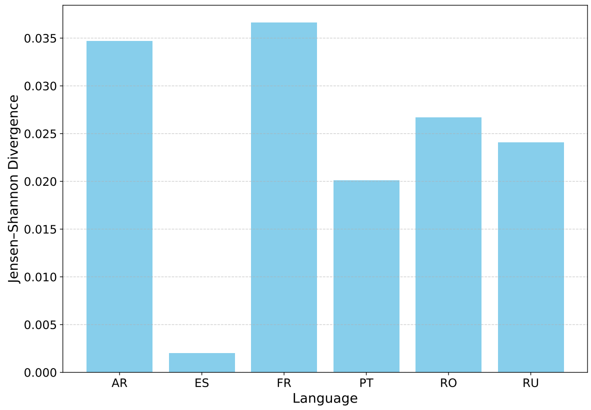
In the topic dimension, all languages display JSD values below 0.16. The highest divergences are observed for French (FR) and Arabic (AR), both near 0.16, followed by Russian (RU) at 0.14, and Portuguese (PT) at approximately 0.13. Spanish (ES) exhibits the lowest topic divergence at just 0.07. These results are consistent with expectations, as economic topics tend to be grounded in technical language that exhibits relatively low cross-lingual variation. The slightly elevated divergence for Arabic and French may reflect differences in syntactic structure or economic reporting style, though the overall divergence remains within a narrow band.

Communication stance shows the lowest overall divergence across dimensions. Even the highest JSD, observed for Arabic (AR) and French (FR), remains at just below 0.04, while all other languages exhibit values below 0.03, with Spanish (ES) approaching 0.00. These findings indicate that temporal markers used to signal forward- versus backward-looking statements are reliably preserved across languages in both translation and classifier representation. This high consistency reinforces the notion that the classification framework is robust in capturing the temporal intent of monetary and financial communication, even across varied linguistic structures.

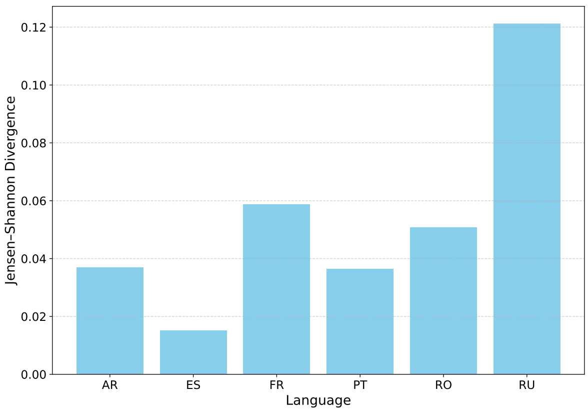
Figure 1: Translation Consistency of Central Bank Communication Across Classification Dimensions



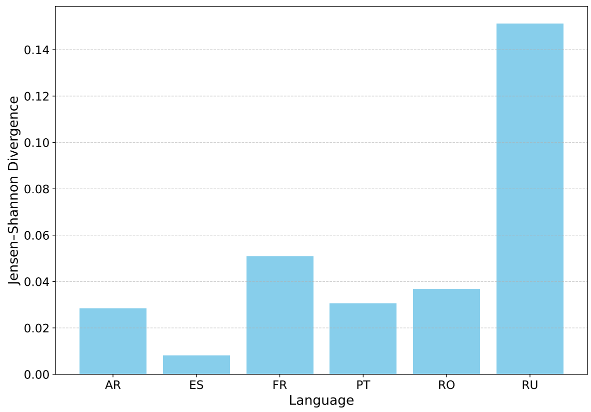
(a) Topic



(b) Communication stance



(c) Audience



(d) Sentiment

Notes: This figure reports the average Jensen–Shannon divergence (JSD) between predicted label distributions for original central bank communications and their translations across multiple languages. JSD measures the semantic distance between two probability distributions. Lower values indicate greater consistency between original and translated texts. Results are shown for four classification dimensions: (a) topic; (b) communication stance; (c) audience; and (d) sentiment.

The audience dimension exhibits slightly more variation, with Russian displaying the highest JSD at 0.12. All other languages fall below 0.06, and Spanish again shows near-perfect consistency (JSD = 0.01). The slightly elevated divergence for Russian may reflect subtle translation artifacts or cross-cultural differences in how institutional actors are referenced. Nevertheless, the overall low values across all languages support the classifier's reliability in identifying intended audiences with minimal distortion.

The sentiment dimension shows a performance similar to the audience's results. The highest JSD is again observed for Russian (0.15), with all other languages remaining below 0.05. Spanish (ES) again exhibits exceptional alignment, with a JSD near 0.01. These findings suggest that, while sentiment is the most linguistically subtle of the four dimensions, the combined translation-classification pipeline still maintains robust consistency in most settings. The elevated divergence for Russian likely reflects both translation challenges and higher classifier uncertainty when detecting affective tone in morphologically rich or less resourced languages.

Taken together, the low divergence values across all languages and dimensions—none exceeding 0.16 and most falling well below 0.10—indicate strong cross-lingual stability in sentence-level classification. These results provide empirical support for using our classifier in multilingual settings. However, slight divergences for specific languages (notably Russian and Arabic) highlight that performance is not entirely invariant across linguistic boundaries and may reflect both translation imperfections and limits in the multilingual generalization of sentence-level embeddings.

Appendix C. Prediction Confidence Analysis

This section examines the distribution of marginal probabilities assigned to predicted labels across the four classification dimensions. Marginal probability is a natural measure of dimension-specific confidence: values near one indicate an unambiguous prediction, whereas lower values reflect either classifier uncertainty or inherent ambiguity in the sentence being classified.

Our classifier estimates joint probabilities over paired dimensions—⟨*topic, communication stance*⟩ and ⟨*audience, sentiment*⟩. However, for confidence analysis, we focus on marginal probabilities derived from these joint outputs. This choice is deliberate: if we were to use the joint probabilities directly, both labels in a pair would receive the same confidence score, masking heterogeneity in classification certainty across individual classes. By contrast, marginal probabilities allow us to isolate how confident the model is in assigning a specific label within a single dimension.

Under this decomposition, the marginal probability of a single label conditioned on x can be computed by summing over the relevant joint distribution and marginalizing out the other coupled dimension. For example, the probability of a specific topic label $y_{\text{topic}} = t$ given a sentence embedding x is:

$$P(y_{\text{topic}} = t \mid x) = \sum_{c \in C} P(y_{\text{topic}} = t, y_{\text{stance}} = c \mid x), \tag{16}$$

where C denotes the set of communication stance labels. Similarly, the probability of a sentiment label $y_{\text{sentiment}} = s$ is:

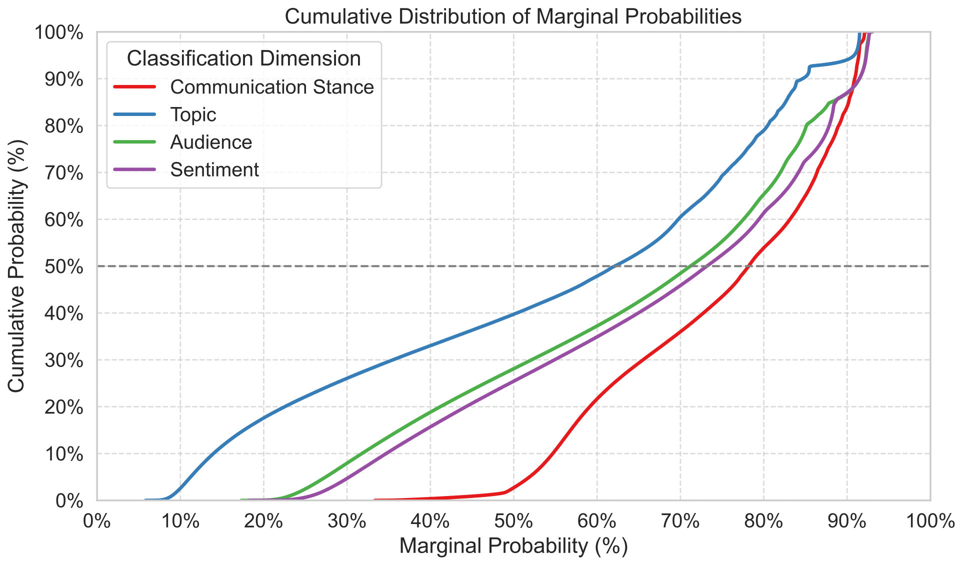
$$P(y_{\text{sentiment}} = s \mid x) = \sum_{a \in \mathcal{A}} P(y_{\text{audience}} = a, y_{\text{sentiment}} = s \mid x), \tag{17}$$

where $\mathcal{A}$ is the set of audience labels.

Figure 1 shows the cumulative distribution of marginal probabilities for each classification dimension. Since the horizontal axis corresponds to the marginal probability and the vertical axis to the cumulative share of classified sentences, the figure directly indicates how often predictions exceed particular confidence thresholds. The results reveal that more than half of all sentences are classified with marginal probabilities above 60 percent for topic, 70 percent for audience and sentiment, and nearly 80 percent for communication stance. The classifier, therefore, assigns relatively high confidence to a large share of predictions, although with clear differences across dimensions. The topic dimension is more challenging to classify, likely attributable to the larger number of topic categories relative to the other classification dimensions.

While the cumulative distributions provide a broad perspective on the classifier's overall confidence across dimensions, they may overlook important heterogeneity at the level of individual classification labels. To uncover these differences, Figure 2 presents boxplots of marginal probabilities for each class within the topic, communication stance, audience, and sentiment dimensions. These plots reveal substantial variation in classification confidence both across and within dimensions. As expected, "Metadata" classifications consistently exhibit the highest marginal probabilities, reflecting the model's ability to identify non-substantive content easily. Among substantive classes, those associated with clearer and more distinct semantic

Figure 1: Predictive Confidence Distribution of Predicted Labels Across Classification Dimensions



Notes: This figure shows the cumulative proportion of sentences with predicted marginal probabilities above a given threshold for each classification dimension. Curves shifted to the right indicate stronger confidence in the classification. For instance, approximately 80 percent of sentences classified by topic receive marginal probabilities greater than 80 percent.

categories—such as ”MP - inflation” in topic, “forward-looking” in communication stance, “financial sector” in audience, and “risk-highlighting” in sentiment—tend to receive higher marginal probabilities. Conversely, labels covering more ambiguous or overlapping content show wider distributions and lower medians. This pattern is particularly evident in the topic dimension, where categories such as “fiscal policy” and “financial stability” exhibit broader probability ranges. At the same time, “supervision and regulation” has the lowest median, consistent with their greater conceptual complexity and potential for overlap.

Figure 2: Distribution of the Classifier's Predictive Probability (Confidence) for Topics

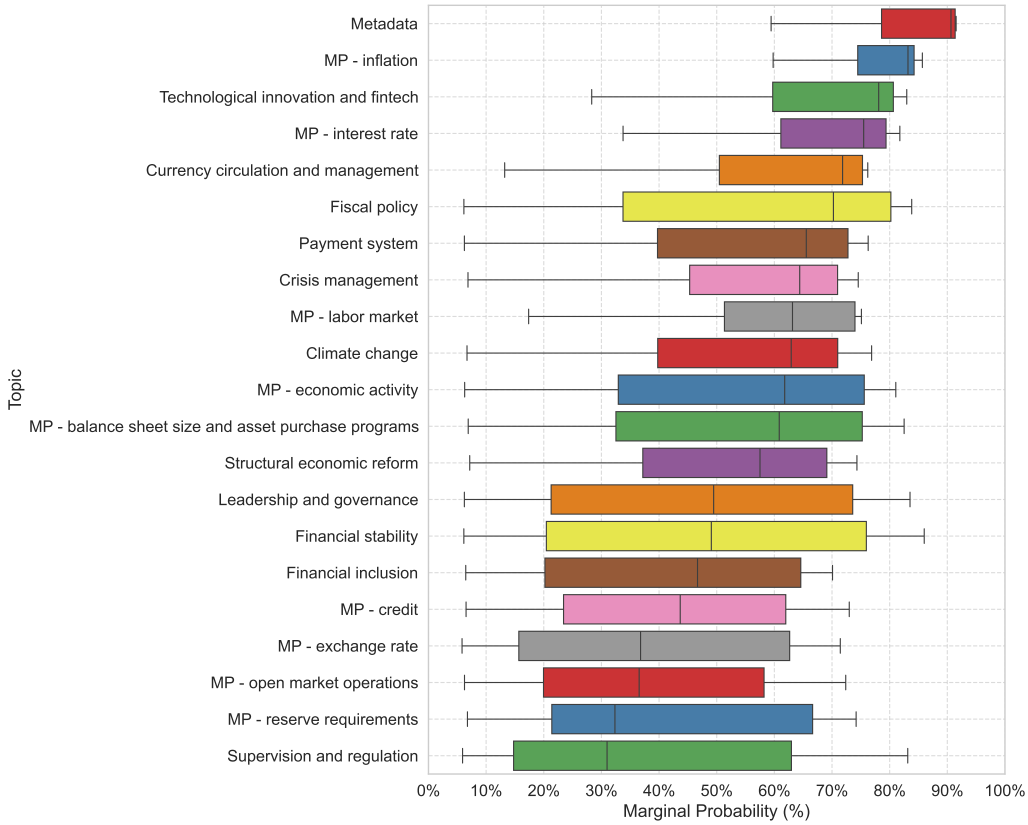


Figure 2 (continued): Distribution of the Classifier's Predictive Probability (Confidence) for Communication Stance

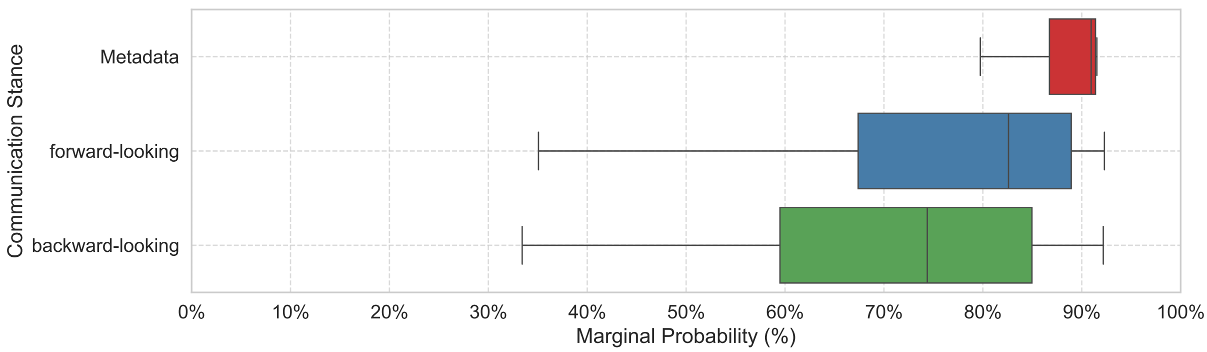


Figure 2 (continued): Distribution of the Classifier's Predictive Probability (Confidence) for Audience

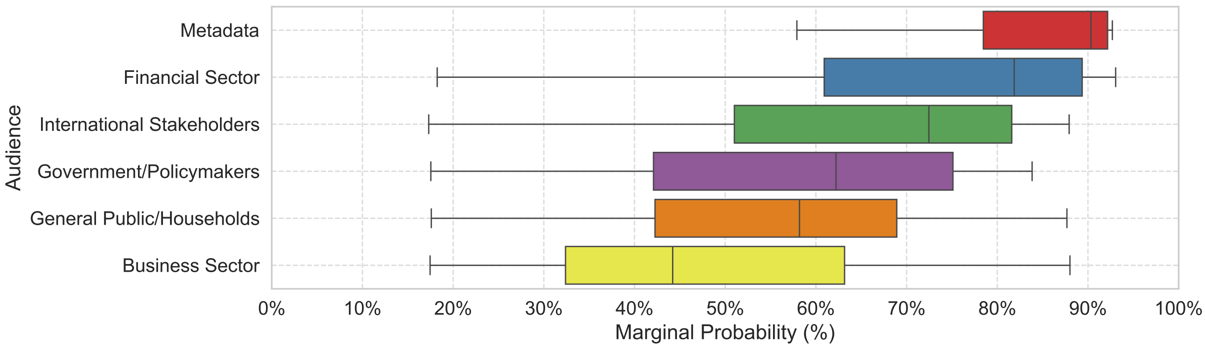
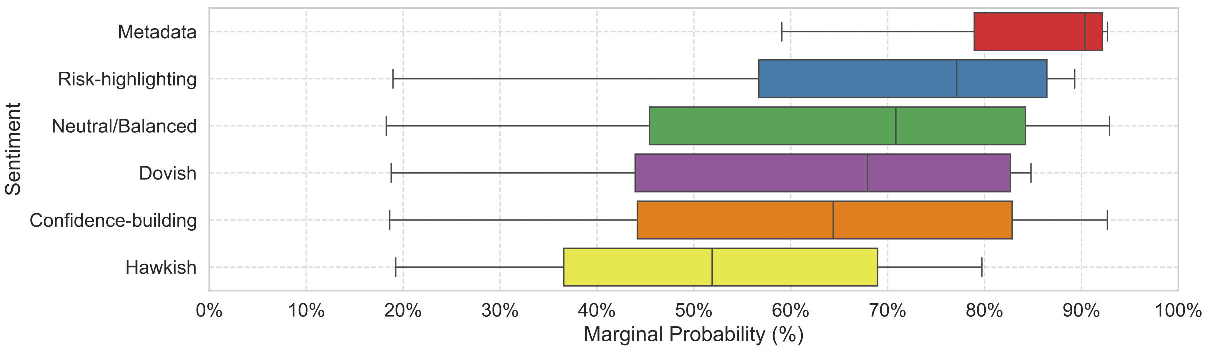


Figure 2 (continued): Distribution of the Classifier's Predictive Probability (Confidence) for Sentiment



Notes: These figures display the distribution of predicted marginal probabilities for each classification label within the four dimensions: topic, communication stance, audience, and sentiment. Each boxplot summarizes the distribution of the classification confidence (classification probability) across all sentences of a specific class. Classes ordered by their median probability. Higher marginal probabilities indicate greater predictive confidence by the classifier in label assignment.

Appendix D. Multilabel Sentences

Central bank communication may address multiple topics and audiences within a single statement, and individual sentences may convey multiple sentiments. Such multilabel sentences present a natural challenge for automated classification models, which assign probabilistic labels reflecting the most salient categories in each dimension. When sentences simultaneously reflect multiple relevant class labels, the classifier distributes probabilities across categories, diluting the marginal probability assigned to any single label. This diffusion lowers the maximum predicted probability and confounds its interpretation as a measure of classification confidence.

To assess the empirical relevance of this issue, we calculate the frequency with which two distinct classes within the same dimension are simultaneously assigned predicted marginal probabilities above 25 percent for the same sentence. This analysis covers three classification dimensions—topic, audience, and sentiment (communication stance is excluded given its binary structure)—and the results are summarized in Figure 1.

Overall, co-occurrence rates are modest, suggesting that multidimensional sentences are not pervasive enough to undermine the interpretability of predicted probabilities. Topic co-occurrence is limited, with most pairs appearing together in fewer than 0.04 percent of sentences. Audience co-occurrence is slightly more common, particularly between the financial and business sectors (1.89 percent), reflecting the natural overlap in messaging toward market participants. Sentiment co-occurrence is somewhat more pronounced, notably between neutral/balanced with confidence-building (3.62 percent) and neutral/balanced with risk-highlighting (3.14 percent), underscoring that central banks often blend supportive and cautionary tones with numerical facts or statements in the same sentence. Nonetheless, these patterns remain sufficiently rare that they do not represent a first-order concern.

An alternative strategy to handle multi-dimensional sentences would involve adopting multilabel classification models, which allow each sentence to be simultaneously assigned to multiple class labels. While such models could more explicitly capture the multi-dimensional nature of certain sentences, they also entail significant practical costs. In particular, multilabel classification requires substantially more complex and labor-intensive annotation, as annotators must identify all applicable categories rather than selecting a single most-relevant label. This raises the threshold for assembling a sufficiently large and representative labeled dataset. In addition, multilabel models often necessitate additional calibration and thresholding choices, which can complicate the interpretation of output probabilities. Given these considerations, and because of the relatively limited empirical relevance of multi-topic sentences in the corpus, the single-label probabilistic approach adopted here provides a pragmatic and analytically robust solution.

Figure 1: Potential Multilabel Co-occurrence Across Sentences – topic dimension

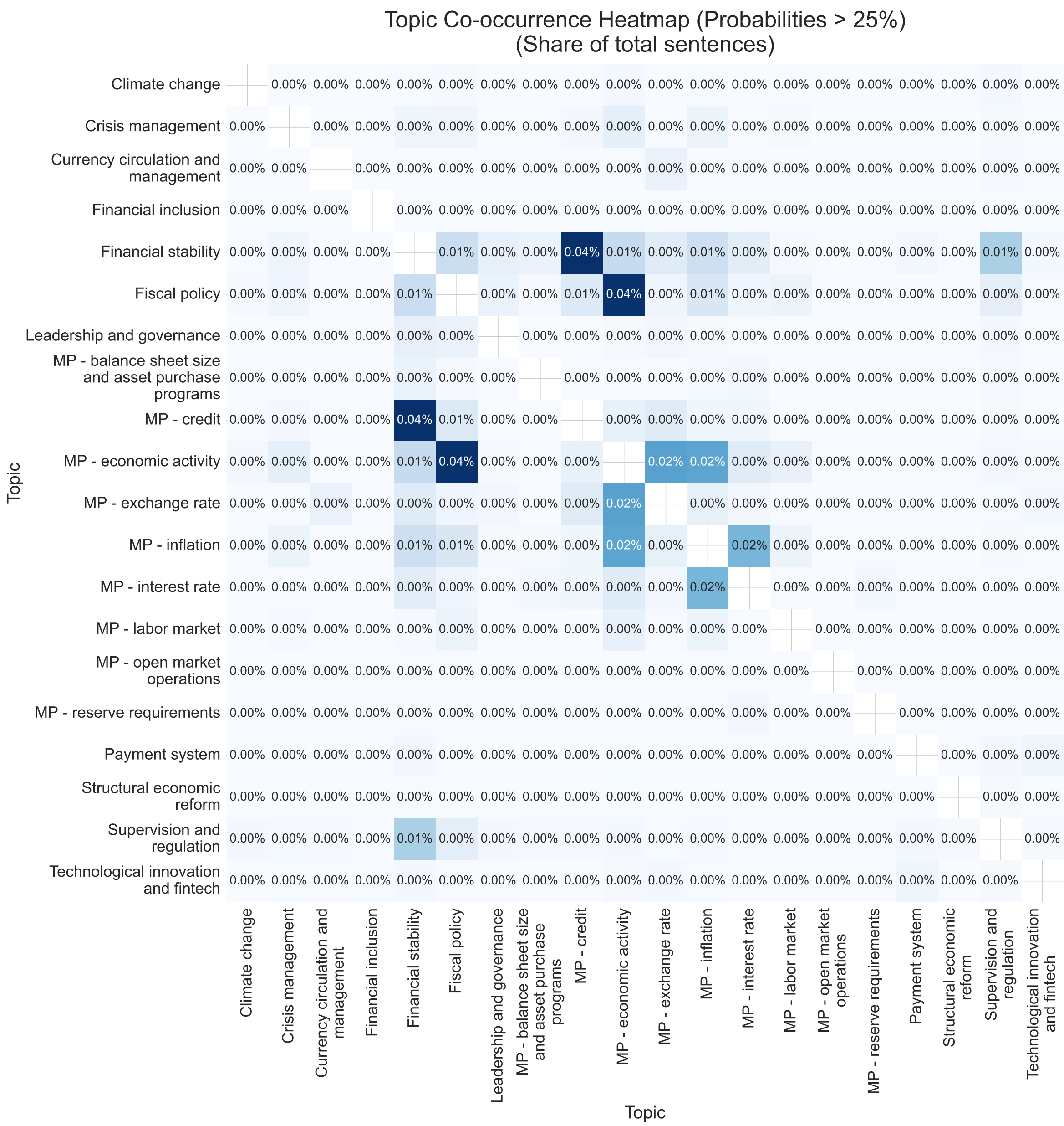


Figure 1 (continued): Potential Multilabel Co-occurrence Across Sentences – audience dimension

Figure 1 (continued): Potential Multilabel Co-occurrence Across Sentences – sentiment dimension

Notes: These figures display co-occurrence heatmaps with the Share of total sentences in which two topics are simultaneously assigned predicted marginal probabilities above 25 percent. Values are expressed in percentage terms, and the main diagonal is excluded to focus on cross-topic associations. Higher co-occurrence rates indicate instances where the classifier identified the presence of multiple relevant topics within the same sentence.

Appendix E. Additional Information

Table 1: Summary statistics of variables used in the regression analysis of the net policy sentiment metric.

Variable	N	Mean	Std. Dev.	Min	P25	P50	P75	Max
Dependent Variables:								
Policy Rate$_{i,t}$	5748	5.69	6.36	-0.75	1.75	4.50	7.50	118.00
Δ Policy Rate$_{i,t+1}$	5735	0.00	0.87	-33.00	0.00	0.00	0.00	20.00
$\lvert\Delta$T-Bill Rate$_{i,t+1}\rvert$	5735	0.23	0.84	0.00	0.00	0.00	0.25	33.00
Δ T-Bill Rate$_{i,t+1}$	4640	-0.03	1.25	-13.55	-0.17	0.00	0.13	16.62
Δ T-Bond Rate$_{i,t+1}$	3361	-0.01	0.68	-12.26	-0.16	0.00	0.14	7.00
Independent Variables:								
Net Policy Sentiment$_{i,t}$	7059	-0.32	0.56	-1.00	-0.75	-0.33	0.00	1.00
Net Policy Sentiment (Fwd)$_{i,t}$	7059	-0.11	0.47	-1.00	-0.33	0.00	0.00	1.00
Net Policy Sentiment (Bwd)$_{i,t}$	7059	-0.18	0.37	-1.00	-0.40	-0.08	0.00	1.00
Straightforwardness Index$_{i,t}$	7059	0.67	0.23	0.00	0.53	0.67	0.82	1.00
Explanation Index$_{i,t}$	7059	9.11	7.32	0.00	5.33	7.89	11.47	120.00
Net Confidence Index$_{i,t}$	7059	-0.22	0.29	-1.00	-0.37	-0.23	-0.08	1.00
Exchange Rate (USD/Local)$_{i,t}$	7054	0.28	0.40	0.00	0.01	0.08	0.36	2.07
CPI (Index)$_{i,t}$	6599	135.04	71.96	32.17	100.71	118.06	146.62	1162.49

Notes: Summary statistics of variables employed in the panel regressions in Section 5.2.1. Variables are expressed in level or change form as indicated. The panel regressions are estimated with monthly data, but the original variation of textual indicators varies by publication schedule for each economy. The net policy sentiment, straightforwardness index, and explanation index are constructed from monetary policy decisions, whose meetings typically occur one to eight times per year, depending on the country. To align with the monthly panel structure, these metrics are forward-filled within their natural frequency interval (e.g., decisions made quarterly are carried forward within the quarter). The net confidence index is computed from a broader set of documents, including annual reports, monetary policy reports, financial stability reports, and monetary policy decisions, and is similarly forward-filled within appropriate intervals. Macroeconomic variables (CPI and exchange rates) are monthly series sourced from the IMF International Financial Statistics and refer to end-of-month values.

Table 2: Summary statistics of variables used in the Granger causality analysis.

	Forward-Looking Confidence	Backward-Looking Confidence	VIX
Observations	420	420	420
Mean	-0.03	-0.04	0.00
Std. Dev.	0.01	0.02	0.02
Minimum	-0.08	-0.13	-0.09
25^{th} Percentile	-0.04	-0.05	-0.01
Median	-0.03	-0.04	0.00
75^{th} Percentile	-0.02	-0.03	0.01
Maximum	0.01	0.04	0.02

Notes: The table summarizes the global monthly series used in the Granger causality analysis used in Section 5.2.4. The forward-looking and backward-looking components of the Net Confidence Index are computed as GDP-weighted averages across countries. VIX is standardized to have a mean of zero and unit variance over the sample period (1990–2025).

References

Ahn, S. C., & Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. *Journal of Econometrics*, *68*, 5–27.

Apel, M., Blix Grimaldi, M., & Hull, I. (2022). How much information do monetary policy committees disclose? Evidence from the FOMC's minutes and transcripts. *Journal of Money, Credit and Banking*, *54*, 1459–1490.

Aruoba, S. B., & Drechsel, T. (2024). *Identifying Monetary Policy Shocks: A Natural Language Approach*. Working Paper 32417 National Bureau of Economic Research.

Ben-David, S., Blitzer, J., Crammer, K., & Pereira, F. (2010). A theory of learning from different domains. In *Machine Learning* (pp. 151–175). Springer volume 79.

Bholat, D., Hansen, S., Santos, P., & Schonhardt-Bailey, C. (2015). *Text Mining for Central Banks*. Technical Report No. 33 Bank of England, Centre for Central Banking Studies.

Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., & Jansen, D.-J. (2008). Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature*, *46*, 910–45.

Born, B., Ehrmann, M., & Fratzscher, M. (2014). Central bank communication on financial stability. *The Economic Journal*, *124*, 701–734.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A. et al. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, *33*, 1877–1901.

Bucher, M. J. J., & Martini, M. (2024). Fine-tuned 'small' LLMs (still) significantly outperform zero-shot generative AI models in text classification, . doi:10.48550/arXiv.2406.08660.

Bundick, B., & Smith, A. L. (2020). The dynamic effects of forward guidance shocks. *Review of Economics and Statistics*, *102*, 946–965.

Campbell, J. R., Evans, C. L., Fisher, J. D., Justiniano, A., Calomiris, C. W., & Woodford, M. (2012). Macroeconomic effects of federal reserve forward guidance. *Brookings Papers on Economic Activity*, (pp. 1–80).

Campiglio, E., Deyris, J., Romelli, D., & Scalisi, G. (2025). *Warning words in a warming world: Central bank communication and climate change*. 418. London School of Economics and Political Science.

Chen, J., Xiao, S., Zhang, P., Luo, K., Lian, D., & Liu, Z. (2024). BGE M3-Embedding: Multi-lingual, multi-functionality, multi-granularity text embeddings through self-knowledge distillation. arXiv:2402.03216.

Cieslak, A., Hansen, S., McMahon, M., & Xiao, S. (2023). *Policymakers' uncertainty*. Technical Report 31849 National Bureau of Economic Research.

Cieslak, A., & Schrimpf, A. (2019). Non-monetary news in central bank communication. *Journal of International Economics*, *118*, 293–315.

Coenen, G., Ehrmann, M., Gaballo, G., Hoffmann, P., Nakov, A., Nardelli, S., Persson, E., & Strasser, G. (2017). *Communication of monetary policy in unconventional times*. 2080. ECB Working Paper Series.

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 8440–8451). Online: Association for Computational Linguistics.

Correa, R., Garud, K., Londono, J. M., & Mislang, N. (2020). Sentiment in central banks' financial stability reports. *Review of Finance*, *25*, 85–120.

de Araujo, D. K. G., Bokan, N., Comazzi, F. A., & Lenza, M. (2025). Word2Prices: embedding central bank communications for inflation prediction. *BIS Working Papers*, *1253*.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In J. Burstein, C. Doran, & T. Solorio (Eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186). Minneapolis, Minnesota: Association for Computational Linguistics.

Ehrmann, M., & Fratzscher, M. (2007). Transparency, disclosure and the Federal Reserve. *International Journal of Central Banking*, *8*, 179–225.

Ehrmann, M., & Talmi, J. (2020). Starting from a blank page? Semantic similarity in central bank communication and market volatility. *Journal of Monetary Economics*, *111*, 48–62.

Gambacorta, L., Kwon, B., Park, T., Patelli, P., & Zhu, S. (2024). CB-LMs: language models for central banking. *BIS Working Papers*, *1215*.

Gertler, M., & Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, *7*, 44–76.

Gorodnichenko, Y., Pham, T., & Talavera, O. (2024). Central bank communication on social media: What, to whom, and how? *Journal of Econometrics*, (p. 105869).

Gürkaynak, R. S., Sack, B. P., & Swanson, E. T. (2005). Do actions speak louder than words? The response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*, *1*, 55–93.

Haldane, A. G., & McMahon, M. (2018). Central bank communications and the general public. *American Economic Review*, *108*, 578–83.

Hansen, A. L., & Kazinnik, S. (2024). Can ChatGPT decipher Fedspeak? *Available at SSRN*, . doi:10.2139/ssrn.4399406.

Hansen, S., McMahon, M., & Prat, A. (2017). Transparency and deliberation within the FOMC: A computational linguistics approach. *Quarterly Journal of Economics*, *133*, 801–870.

Hansen, S., McMahon, M., & Tong, M. (2019). The long-run information effect of central bank communication. *Journal of Monetary Economics*, *108*, 185–202.

Hendrycks, D., Liu, X., Lee, M., Mazeika, M., Song, D., & Steinhardt, J. (2021). The many faces of robustness: A critical analysis of out-of-distribution generalization. In *International Conference on Computer Vision (ICCV)* (pp. 8340–8349).

Hubert, P., & Labondance, F. (2021). The signaling effects of central bank tone. *European Economic Review*, *133*, 103684.

International Monetary Fund (2025). Annual Report on Exchange Arrangements and Exchange Restrictions (AREAR). URL: https://www.elibrary-areaer.imf.org/Pages/Home.aspx available from the International Monetary Fund.

Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Maschinot, A., Liu, C., & Krishnan, D. (2020). Supervised contrastive learning. *Advances in Neural Information Processing Systems*,

33, 18661–18673.
Lin, J. (1991). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, *37*, 145–151.
Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, *66*, 35–65.
van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579–2605.
Nakamura, E., & Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: The information effect. *Quarterly Journal of Economics*, *133*, 1283–1330.
Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, (pp. 1417–1426).
Pfeifer, M., & Marohl, V. P. (2023). CentralBankRoBERTa: A fine-tuned large language model for central bank communications. *Journal of Finance and Data Science*, *9*, 100114.
Recht, B., Roelofs, R., Schmidt, L., & Shankar, V. (2019). Do ImageNet classifiers generalize to ImageNet? In *International Conference on Machine Learning (ICML)* (pp. 5389–5400). PMLR.
Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In K. Inui, J. Jiang, V. Ng, & X. Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3982–3992). Hong Kong, China: Association for Computational Linguistics.
Romer, C. D., & Romer, D. H. (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review*, *94*, 1055–1084.
Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2022). Measuring news sentiment. *Journal of Econometrics*, *228*, 221–243.
Siddhant, A., Hu, J., Johnson, M., Firat, O., & Ruder, S. (2020). Xtreme: A massively multilingual multi-task benchmark for evaluating cross-lingual generalization. In *Proceedings of the International Conference on Machine Learning 2020* (pp. 4411–4421).
Silva, T. C., & Zhao, L. (2016). *Machine learning in complex networks*. Springer.
Sturm, J.-E., & De Haan, J. (2011). Does central bank communication really lead to better forecasts of policy decisions? New evidence based on a Taylor rule model for the ECB. *Review of World Economics*, *147*, 41–58.
Swanson, E. T. (2021). Measuring the effects of Federal Reserve forward guidance and asset purchases on financial markets. *Journal of Monetary Economics*, *118*, 32–53.
Tunstall, L., Reimers, N., Jo, U. E. S., Bates, L., Korat, D., Wasserblat, M., & Pereg, O. (2022). Efficient few-shot learning without prompts. doi:10.48550/ARXIV.2209.11055.
Woodford, M. (2005). *Central Bank Communication and Policy Effectiveness*. Working Paper 11898 National Bureau of Economic Research.

PUBLICATIONS

From Text to Quantified Insights: A Large-Scale LLM Analysis of Central Bank Communication

Working Paper No. WP/2025/109