IFC-Bank of Italy Workshop on "Data science in central banking"

18-20 February 2025

The monetary policy statement database[1]

Cory Baird,
University of Tokyo

Jonathan Benchimol,
Bank of Israel

Vira Vyshnevska and Wook Sohn,
KDI School

Iegor Vyshnevskyi,
Sogang University

[1] This contribution was prepared for the workshop. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.

The Monetary Policy Statement Database: An LLM Application to Global Financial Conditions*

Cory Baird† Jonathan Benchimol‡ Wook Sohn§ Vira Vyshnevska§
Iegor Vyshnevskyi¶

March 11, 2026

Abstract

This study introduces the Monetary Policy Statement Database (MPSD), comprising 6,693 statements from 51 central banks worldwide (1990–2024). We develop a reproducible pipeline combining standard natural language preprocessing with large language model (LLM) tools for cross-country analysis. Four key findings emerge. First, statements lengthened substantially after the Global Financial Crisis while readability improved modestly. Second, inflation references comove across countries during global inflation episodes. Third, LLM-based question answering and aspect-based sentiment reveal that central banks attribute global financial conditions primarily to broad U.S. macroeconomic developments rather than to Federal Reserve policy actions specifically. Fourth, using a benchmark dictionary-based sentiment index and LLM-derived aspect-based sentiment indicators, Granger causality tests suggest that statement sentiment predicts the Global Financial Cycle rather than merely responding to it. The MPSD and accompanying codebase support reproducible research on monetary policy communication and international transmission.

Keywords: Central bank communication, Large language models, Text analysis, Generative database, Machine learning.

JEL Classification: C55, C63, E52, E58, G15.

*This paper does not necessarily reflect the views of the Bank of Israel or any other institution.
†University of Tokyo, Tokyo, Japan. Corresponding author. Email: cbaird@g.ecc.u-tokyo.ac.jp
‡Bank of Israel, Jerusalem, Israel.
§KDI School, Sejong, South Korea.
¶Sogang University, Seoul, South Korea.

1 Introduction

The evolution of monetary policy over recent decades has elevated central bank communication to a primary policy instrument, fundamentally transforming how monetary authorities influence economic outcomes (Woodford, 2005; Blinder, 2018). As former Federal Reserve Chairman Ben Bernanke observed, monetary policy has become "98 percent talk and 2 percent action," reflecting the critical role of communication in shaping market expectations and enhancing policy effectiveness (Blinder et al., 2008). This transformation has become particularly salient as central banks increasingly rely on forward guidance and other communication-based tools to influence both financial markets and household expectations (Hansen et al., 2019; Coibion et al., 2022).

The empirical analysis of central bank communications faces several critical constraints that limit our understanding of how monetary policy messages are crafted, transmitted, and interpreted across global markets. First, despite significant advances in communication strategies (de Haan and Hoogduin, 2024), the absence of a centralized, standardized database of monetary policy communications has restricted analysis to individual institutions (Born et al., 2014; Benchimol et al., 2025) and small groups of advanced or emerging economies (Cieslak and Schrimpf, 2019; Vyshnevskyi et al., 2024). Current research relies primarily on the Bank for International Settlements' repository of speeches, leaving a significant gap in the systematic analysis of monetary policy statements (Picault and Renault, 2017). Second, the lack of robust, open-source preprocessing pipelines has hindered the reproducibility of text-based monetary policy research, a methodological concern highlighted by Gentzkow et al. (2019) and Hansen and McMahon (2016). Third, while natural language processing (NLP) techniques have advanced rapidly, their application to monetary policy analysis has been constrained by insufficient computational infrastructure and domain-specific training data.

Recent studies confirm the complex and crucial influence of central bank communication. Hansen and McMahon (2016) demonstrate how linguistic analysis effectively reveals policy intentions and economic forecasts. Kryvtsov and Petersen (2021) provide experimental evidence on the effectiveness of different communication strategies, while Hayo and Neuenkirch (2015) document communication's crucial role during financial crises. These studies, alongside research on information effects by Nakamura and Steinsson (2018) and Bauer and Swanson (2023), highlight both the potential and challenges of systematically analyzing monetary policy communications.

This paper addresses these challenges by introducing the Monetary Policy Statement Database (MPSD), a comprehensive infrastructure for analyzing monetary policy statements across 51 central banks. Our contribution is fourfold.

First, we develop and publicly release a standardized database of monetary policy statements enabling rigorous cross-country analysis of central bank communications. This resource empowers researchers to investigate policy transmission and coordination across diverse economies, including those beyond the traditional focus on major central banks (Armelius et al., 2020).

Second, we implement a sophisticated technical architecture that combines advanced NLP capabilities with robust code and data management tools. The infrastructure leverages data pipeline management (DVC), version control (Git), and experiment tracking while integrating cloud computing solutions through Amazon S3 buckets for storage and Hugging Face for open-source language model integration. This framework supports automated text summarization and question answering (QA) systems specifically designed for central bank applications, enabling efficient analysis of large-scale monetary policy communications. Following best practices in computational reproducibility (Gentzkow et al., 2019), our implementation uses object-oriented programming principles, with modular Python classes providing standardized tools for text preprocessing and advanced NLP analytics.

Third, we offer a practical framework enabling users to leverage large language models (LLMs), including Claude, Gemini, and ChatGPT. Building on methodological advances in text analysis, we develop tools for extracting policy-relevant information from complex monetary policy documents (Gonzalez and Tadle, 2022).

Fourth, we provide novel empirical evidence on the communication channel of international policy transmission. Our analysis yields four key findings with implications for understanding global monetary policy coordination: (i) Central banks worldwide exhibit remarkable narrative convergence during major economic shocks, with communications showing strong thematic comovement that precedes—rather than follows—observable financial market stress. (ii) When attributing global financial conditions to external factors, central banks systematically cite broad U.S. economic developments (housing markets, consumer spending, growth prospects) rather than Federal Reserve policy actions specifically. This distinction has important implications for modeling policy spillovers. (iii) Dictionary-based communication sentiment and LLM-derived sentiment indicators Granger-cause movements in the Global Financial Cycle (GFC), suggesting that central bank narratives actively shape global financial conditions rather than merely describing them. (iv) Substantial heterogeneity exists in how advanced-economy versus emerging-market central banks discuss global conditions, with emerging markets showing more idiosyncratic patterns, particularly at times of financial stress. These findings suggest that communication may be an underappreciated channel through which monetary policy transmits across borders.

Table 1 illustrates our approach to ensuring end-to-end transparency in artificial intelligence (AI) applications for monetary policy analysis. The pipeline shows how we maintain scientific rigor from raw data collection through preprocessing, model inference, and analysis—addressing concerns about reproducibility in computational economic research, particularly when applying advanced language models to policy analysis.

Table 1: End-to-End Transparency for Replicable AI

Data Collection	**Featurization**	**Replicable AI Use**
Identify statement links	Cleaning	Replicable LLM infrastructure
Scrape statements	Sentence segmentation	Track inputs (prompts, texts)
Automate retrieval	Part-of-speech tagging	Track outputs (responses)

Source: Authors.

We leverage our data and AI infrastructure to investigate several key questions regarding monetary policy, the Global Financial Cycle (Miranda-Agrippino and Rey, 2020; Rey, 2015), and the transmission of policy narratives across borders (Armelius et al., 2020). We employ LLMs for QA and aspect-based sentiment (ABS) analysis. In QA, we prompt the model to summarize the drivers of the global economy, particularly focusing on instances where central banks identify the Federal Reserve as a key influence. For ABS, we use LLMs to extract sentiment for specific themes related to the impact of global financial conditions on domestic economies, such as economic growth, financial conditions, capital flows, and monetary policy spillovers.

Our findings from this LLM analysis reveal that central banks often cite the U.S. economy as a driver of global economic conditions rather than specifically attributing influence to the Federal Reserve. This nuanced distinction—where broader economic factors in the United States are perceived as more impactful than central bank actions alone—has important implications for modeling policy spillovers. Furthermore, our sentiment analysis shows that central banks expressed greater concern about growth and financial conditions than monetary policy and capital flows in the post-2008 period, likely reflecting the lasting impact of the Global Financial Crisis and increased global economic interconnectedness.

While recent discussions of AI in economics have focused primarily on closed-source foundation models, we emphasize the fundamental importance of robust, open-source data science infrastructure. The MPSD provides comprehensive tools for text analysis specifically designed for central bank communications, enabling researchers to extract and analyze policy narratives systematically across institutions and time. This approach builds on seminal work in computational text analysis (Gentzkow et al., 2019) while addressing specific challenges in monetary policy research (Hansen and McMahon, 2016).

Our work connects to the growing literature on international monetary policy spillovers. Obstfeld (2015) highlights the challenges that global financial integration poses for monetary policy independence, Bruno and Shin (2015) demonstrate significant effects of U.S. monetary policy on capital flows to emerging markets, and Kalemli-Özcan (2019) show that these spillovers primarily operate through global banks and risk premia in international asset markets. We extend this literature by focusing specifically on the communication channel of policy spillovers, providing direct evidence that central bank narratives—beyond policy actions themselves—significantly influence global financial conditions. Our findings complement Brusa et al. (2020), who document that ECB communications affect U.S. asset prices through signaling about global economic conditions.

Our work contributes to multiple strands of literature. First, we extend research applying NLP to central bank communications (Gonzalez and Tadle, 2022), moving beyond specific aspects like sentiment (Loughran and McDonald, 2011; Correa et al., 2021) or clarity (Bulír et al., 2013; Benchimol et al., 2023) to provide a comprehensive analytical framework. Second, we advance the literature on monetary policy transmission by enabling systematic cross-country analysis of communication strategies and their effects (Hansen et al., 2019; Picault and Renault, 2017). Third, we contribute to the literature on global financial cycles and policy spillovers (Miranda-Agrippino and Rey, 2020; Armelius et al., 2020). Fourth, we extend the methodological literature on text analysis in economics (Gentzkow et al., 2019; Benchimol et al., 2022) by developing reproducible procedures specifically for central bank communications. This project also complements recent large-scale data collection efforts in central bank communication research. In particular, Campiglio et al. (2025) introduce the cbspeeches.com corpus comprising over 35,000 speeches from 131 central banks (1986–2023). While their corpus focuses on speeches and covers a broader institutional range, our database concentrates specifically on monetary policy statements—the official communications released in direct connection with policy rate decisions. The two resources are therefore complementary and together support a comprehensive ecosystem for reproducible research on central bank communication.

Our approach also connects to recent surveys on central bank communication. Masciandaro et al. (2024) provide a comprehensive review of how central banks have evolved from silence to active social media engagement, documenting techniques used to analyze communication effectiveness. Our infrastructure complements their focus on social media by providing systematic analysis of formal policy statements. Additionally, our work relates to the broader literature on monetary policy transmission mechanisms. De Grauwe and Ji (2023) demonstrate that using forward-looking data in monetary policy affects policymaking quality differently across policy regimes, while Kumar et al. (2023) show that market volatility significantly influences the term premium through channels interacting with monetary policy. These findings underscore the importance of understanding how central bank communications shape expectations and market responses—precisely the mechanisms our database enables researchers to investigate.

The remainder of the paper is organized as follows. Section 2 details the technical architecture and construction of the MPSD, including data collection procedures, preprocessing pipelines, and the LLM inference framework. Section 3 presents an exploratory analysis of the database, highlighting key patterns in central bank communications across time and institutions. Section 4 examines our application of generative AI to central bank communications, focusing on extracting economic narratives related to the GFC. Section 5 discusses policy implications and potential extensions, and Section 6 concludes.

2 Database Construction and Architecture

The MPSD constitutes a comprehensive research infrastructure designed to facilitate systematic analysis of monetary policy communications across global central banks. The database encompasses monetary policy statements from 51 central banks, with coverage varying by institution. Statement frequency ranges from monthly to quarterly depending on each central bank's policy framework, yielding 6,693 unique policy statements.

We define a monetary policy statement observation as a text communication released in direct connection with a policy rate decision event (whether a change or no change in rates). Because central banks employ varying naming conventions—including "decisions," "statements," "announcements," and "press releases"—we systematically map each institution's official document series to the corresponding decision dates. This mapping is fully documented in the accompanying metadata file. For example, the ECB's communication evolved from the "Introductory Statement" (1999–mid-2021) to the "Monetary Policy Statement" (since mid-2021), while short "Monetary Policy Decisions" press releases—which historically contained only rate decisions without supporting rationale—are not treated as equivalent to full policy statements in our corpus. Similarly, for the Reserve Bank of Australia, we include the post-meeting "Monetary Policy Decision" statements rather than the separate quarterly "Statement on Monetary Policy" inflation report. The corpus excludes ancillary documents such as inflation reports, minutes, and speeches, which serve different communicative functions. All documents were preprocessed to remove non-textual elements including tables, charts, and lists of meeting attendees, ensuring that linguistic analysis captures substantive policy communication content.

Monetary policy communications take different formal labels across central banks, including "decisions," "statements," and "press releases," as documented in Table 4. While these document types can be identified, they are not treated as separate analytical categories in this study. Despite differences in naming conventions and format, these communications are released in direct connection with policy rate decision events (change or no change) and serve a functionally comparable role: conveying the policy outcome together with varying degrees of economic assessment and forward-looking guidance. In many cases, a single document simultaneously fulfills multiple functions, making a clean and objective partition by document type difficult. To preserve transparency and cross-country comparability, we therefore treat these texts as institution-specific realizations of a common policy-decision communication instrument, while explicitly accounting for observable sources of possible heterogeneity such as language of publication.

The technical architecture implements three fundamental design principles that address specific challenges in monetary policy research. The infrastructure enables scalable data collection and standardization through advanced web scraping and preprocessing pipelines, essential for maintaining consistency across the diverse spectrum of central bank websites and communication formats (PDF, DOC/DOCX, HTML). We also maintain data integrity through version control, addressing reproducibility challenges identified in computational social science research (Gentzkow et al., 2019).

2.1 Technical Infrastructure Implementation

The MPSD ensures integrity of economic research data by strictly enforcing lineage. Using Data Version Control (DVC), the system tracks the full data lifecycle—from web scraping to LLM-based analysis—via parameterized pipelines. To support full computational reproducibility, the architecture couples Git for code versioning with uv for unified Python version and dependency management (ensuring that data manipulation tools such as pandas work consistently across different environments).

Table 2 illustrates the LLM infrastructure we developed specifically for central bank communication analysis. This diagram shows the integration of multiple components (modules), including text retrieval, preprocessing, and LLM integration. The structure enables seamless incorporation of different AI models (such as ChatGPT and Gemini, as well as open-source models hosted on Hugging Face) and supports various analytical tasks including classification, summarization, and QA. This modular design ensures reproducibility while maintaining flexibility for users to adapt the infrastructure to their specific use cases.

Data storage relies on Amazon S3 buckets managed directly through DVC. The system segregates assets into distinct partitions—raw data, processed outputs, and model artifacts—ensuring clear data lineage and secure access control.

2.2 Code and Folder Infrastructure

The MPSD employs a modular design with discrete functional components for text retrieval, text cleaning, and LLM inference. As illustrated in Figure 1, the codebase follows a standardized folder structure based on the "cookiecutter" template for data science projects. This approach organizes the project into logical sections to facilitate maintenance and collaboration:

Table 2: AI Infrastructure for Central Bank Communication Analysis

Modular LLM-based Data Processing Framework for Academic Research		
Data Acquisition Module (Abstract Class)	**LLM Processing Module (Abstract Class)**	**LLM Data Pipeline Workflow**
Methods: 1. Initialize (source parameters) 2. Scrape Statement Links 3. Download/Save Statements	**Methods:** 1. Initialize (LLM Model) 2. Process/Load Prompts 3. Generate Response	**Methods:** 1. Initialize (Input Corpus) 2. Filter via Keywords or POS 3. LLM Inference 4. Save Response
Output: Text Files CSV/Parquet Text Dataframes	**Output:** Response (Text/JSON) Prompt (Text/JSON) Model Parameters (YAML)	**Output:** Response (Text/JSON) Prompt (Text/JSON) Model Parameters (YAML) Figure for Research (PNG/PDF)
Abstraction: Extend abstract functions to accommodate new data sources (e.g., central bank speeches, minutes) and communication modes.	**Abstraction:** Implement model-agnostic functions to leverage diverse LLM architectures (e.g., Gemini, Claude, HuggingFace models).	**Use-Case:** Utilize concise, reusable scripts for seamless integration with LLM APIs (e.g., Gemini, Claude) by researchers.

Source: Authors.

- `src`: Contains Python scripts organized by functionality
- `data`: Stores raw and transformed text files and processed outputs
- `references`: Houses GenAI prompts and documentation

Complementing this structure, the system implements automated database updates through DVC pipelines. This enables researchers to access data and replicate the entire analysis—including the graphs and charts utilized in this paper—with a single command. This structure ensures end-to-end reproducibility from data manipulation through NLP analysis, addressing a critical need in text analysis.

2.3 Data Collection and Processing Framework

Our database houses two data types: statements retrieved programmatically via Python, stored in the 'API' folder, and statements downloaded manually, termed 'static downloads' (Vyshnevskyi et al., 2024). As illustrated in Table 2, the process for programmatically sourced data uses an agnostic class with three abstract methods: (1) downloading statement links, (2) iterating through these links to download each statement, and (3) parsing each statement and saving the output as Parquet and CSV files. The pipeline's final stage applies format-specific processing for PDF, Word, and HTML documents.

Figure 1: Folder Structure Based on "Cookiecutter" Template for Data Science Projects

```
project_root/
├── src/
│   ├── data_processor.py
│   ├── figure_generator.py
│   ├── processor.py
│   ├── pipeline.py
│   ├── sentiment_lexical.py
│   ├── sentence_split.py
│   └── web_scraper.py
├── references/
│   ├── dictionaries/
│   │   ├── centralbank_info.py
│   │   ├── dict_regions.py
│   │   ├── dict_country_converter.py
│   │   └── sentiment_label.py
│   ├── lexical_sentiment_dictionary/
│   │   └── lexical_master_dict.xlsx
│   └── paper_econsurveys/
│       ├── all_figures_settings.yaml
│       ├── llm_sentiment_prompt.yaml
│       └── llm_gfc_driver_prompt.yaml
├── data/
│   └── database/
│       └── txt_files/
├── external/
├── sentences/
└── all_analytical_output/
    ├── cache/
    │   └── [Intermediate processed data (.parquet)]
    ├── tables/
    │   └── [Final manuscript tables (.xlsx, .tex)]
    ├── sentiment/
    │   └── llm_processsor/
    │       └── [LLM sentiment scoring outputs (.csv, .parquet)]
    └── figures/
        └── [15+ Final Analytical Figures (.pdf)]
```

Source: Authors.

The Python code employs a modular structure, maintaining an agnostic design while permitting targeted modifications. Base classes define abstract interfaces (including the three methods previously mentioned), complemented by specific implementation classes tailored to individual central bank websites to handle their variations. We have decided not to release the scraping code due to concerns about potential misuse that could overload central bank web servers. This approach is consistent with established practices in empirical economics, where reproducibility is defined by the ability to replicate published results using shared data and code rather than by replicating the original data collection process itself. We hope our project highlights best practices for similar text database initiatives, demonstrating how abstract and agnostic designs can help optimally organize scraping code.

Continuous integration through GitHub Actions implements sophisticated workflows for database maintenance and updates. These workflows allow us to perform frequent checks for new policy statements, enabling us to cross-check any findings against the links stored in our system.

2.4 LLM Inference Framework

The MPSD incorporates a framework for leveraging language models in central bank applications. We created an abstract LLM class that seamlessly integrates Claude, ChatGPT, Gemini, and open-source models hosted on Hugging Face. This class provides methods to select models, process prompts from YAML files, and generate and save responses, enabling researchers to test various prompts and assess the robustness of their results across multiple models.

2.5 Future Development and Maintenance

The MPSD features a forward-looking architecture designed for scalability to incorporate other types of central bank communication, such as speeches, summaries, and minutes. The modular design also enables easy addition of new central banks. Furthermore, integration with GitHub and DVC enables automated workflows that ensure the database is regularly updated while preserving versioned, point-in-time snapshots of the data.

3 Database Content

3.1 Interest Rate Announcements

Over the last century, central bank communication has evolved from indirect and cryptic messaging to becoming an independent and significant monetary policy tool in both developed and developing economies (Blinder et al., 2024). This communication involves disseminating information about monetary policy objectives, economic outlook, and future policy trajectories (Blinder et al., 2008). Assenmacher et al. (2021) defines monetary policy communication as a central bank's transmission of its objectives, plans, instruments, policy decisions, and economic assessments.

Monetary policy statements—official communications explaining policy decisions and their underlying rationales (Dincer et al., 2022)—represent a crucial component of this communication framework. Their primary objectives include managing inflation expectations, enhancing policy predictability, ensuring transparency, and building institutional credibility with markets, governments, and the public (Kahn, 2007; Blinder et al., 2008; Issing, 2019; Lewis et al., 2020).

Central banks employ various nomenclature for these statements: the Reserve Bank of Australia (RBA) issues a post-meeting "Monetary Policy Decision," the Bank of Japan (BoJ) releases a "Statement on Monetary Policy," the Federal Reserve's primary monetary policymaking body (Federal Open Market Committee) produces a "FOMC Statement," and the European Central Bank (ECB) publishes an "Introductory Statement" (until mid-2021) or "Monetary Policy Statement" (since mid-2021)—documents that provide the detailed rationale accompanying each policy decision. This practice is relatively recent, with the RBA releasing one of the first official policy statements on January 23, 1990, following New Zealand's pioneering efforts in transparent communication during the mid-1990s (Blinder et al., 2001).

While the Australian media reported on the RBA's January 1990 monetary policy change, a notable article in a business-focused newspaper, potentially the first centered on this type of monetary policy statement, did not explicitly describe the statement itself as a new communication format.[1] Searches in other publications yielded similar findings.

The statement's novelty as a communication tool may have been overlooked because senior RBA officials were already actively communicating with markets and the public through other channels, such as speeches, prior to the January 23, 1990, statement, as indicated by news searches from that period.[2] This practice of announcing policy changes began formally in January 1990, marking a shift from previous operational methods where market participants inferred policy changes from the RBA's market actions.

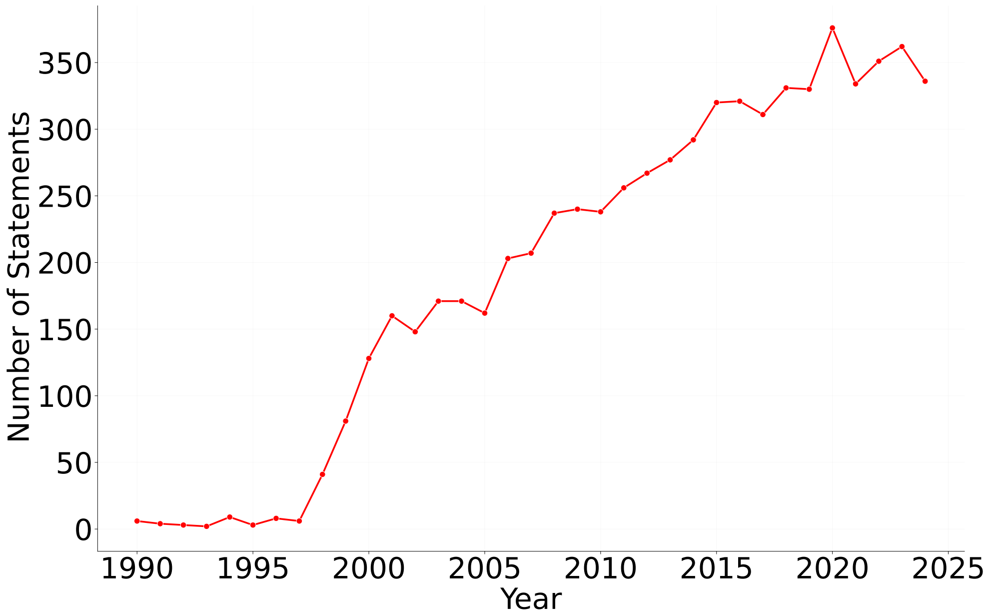
The U.S. Federal Reserve issued its first press release on February 4, 1994, initially publishing statements only when modifying the federal funds rate. By May 1999, the FOMC began releasing statements after every meeting, and by May 2002, included voting details and dissenting vote explanations (Acosta and Meade, 2015). Other central banks followed similar patterns: the Reserve Bank of New Zealand in 1996, the BoJ in 1998, the Swiss National Bank in 2000, and the ECB in 2001. As Figure 2 illustrates, statement issuance accelerated significantly following the 2008 financial crisis.

While formatting varies across institutions, monetary policy statements

[1]Tim Dodd. (January 24, 1990 Wednesday). Timing Surprises Markets. Australian Financial Review.

[2]Tony Kaye. (April 12, 1989 Wednesday). Timing 'Crucial' in Monetary Policy.

Figure 2: Annual Monetary Policy Statements in the Database



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

typically include inflation and growth forecasts, economic assessments, and justifications for policy adjustments (Blinder, 2009). Federal Reserve statements generally assess domestic economic conditions, reference the dual mandate, announce policy decisions, provide forward guidance, and report voting outcomes (Acosta and Meade, 2015). The ECB structures its statements around economic and monetary analyses, policy decisions, rationales, and forward guidance (Assenmacher et al., 2021). The Bank of Korea employs a six-paragraph framework covering policy decisions, global and domestic economic conditions, inflation trends, financial market developments, and future policy direction. These statements typically become more detailed during economic crises, such as the 2008 financial crisis or the COVID-19 pandemic (Macklem and Vardy, 2023).

Statement style has evolved toward greater accessibility and clarity (Blinder et al., 2024), increasingly incorporating visual elements to communicate complex concepts (Assenmacher et al., 2021). Most central banks have converged on approximately eight annual policy meetings, balancing the need for policy flexibility with regular updates. Notable exceptions include the Swiss National Bank (four annually) and the RBA (eleven annually, reduced to eight in 2024). As meeting frequency has declined at some institutions, statement length and detail have increased, particularly for the FOMC (Davis and Wynne, 2019), reflecting the rise of non-standard monetary policy tools such as forward guidance and asset purchases.

The monetary theory consensus now emphasizes open communication as essential for effective policy transmission (Eijffinger and Masciandaro, 2014). Research demonstrates communication's importance in improving monetary transmission and maintaining trust (Ehrmann and Wabitsch, 2022). Global monetary policy has become increasingly transparent, with central banks—including those in emerging economies—publishing policy statements to enhance market stability and policy effectiveness (Dincer et al., 2022). Despite this expanded outreach, audience engagement remains uneven: firms and consumers demonstrate limited awareness, while financial market participants respond significantly to communications, including policy statements (Blinder et al., 2024). Semantic analyses of ECB, Federal Reserve, and other central bank statements reveal a trend toward increased precision and transparency following the Global Financial Crisis (Benchimol et al., 2020).

3.2 Database Overview

This section examines the evolution and characteristics of monetary policy statements (MPS). We define a *monetary policy statement observation* as a written statement released by a central bank on date t, associated with a policy-rate decision event (change or no change). To account for variations in naming conventions across central banks, we map each bank's official document series—released on or explicitly linked to a decision date—to the corresponding decision event t and document this mapping. We begin with an overview of our database, which encompasses approximately 6,693 unique monetary policy statements from 51 central banks globally, spanning the period from 1990 to 2024. As depicted in Figure 8, this standardized and diverse corpus provides an extensive dataset ideal for training and benchmarking AI systems.

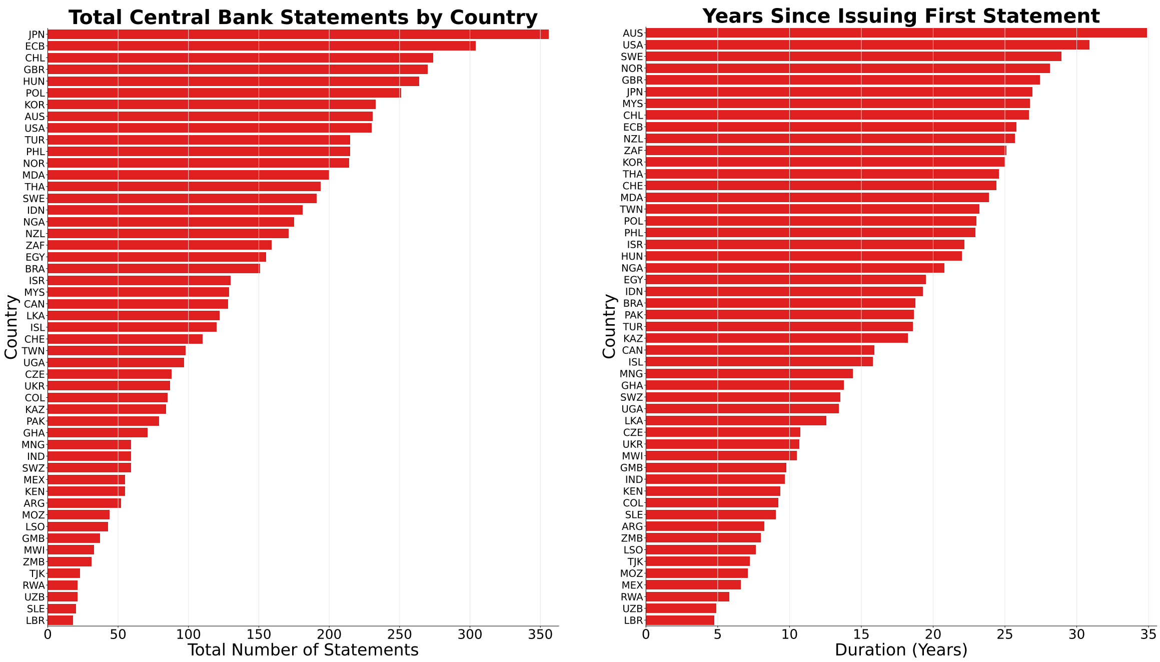
A descriptive analysis of our dataset, as summarized in Tables 3 and 4, reveals substantial heterogeneity in central bank communication patterns across institutions.[3] The Bank of Japan (BoJ) and European Central Bank (ECB) emerge as prolific publishers of statements, while central banks in middle-income countries exhibit more limited publication frequency and shorter historical coverage. These cross-institutional variations reflect differences in institutional frameworks and communication strategies, as illustrated in Figure 3. The statistics on coverage periods, statement frequency, linguistic characteristics, and readability metrics demonstrate systematic differences in communication style and complexity, highlighting the evolution of central bank transparency practices.

Figure 3 provides comparative evidence of central bank communication practices across countries, with Panel (a) documenting the total number of monetary policy statements per country and Panel (b) illustrating the duration of coverage in years. The data reveal considerable heterogeneity in both dimensions. For instance, while the BoJ leads in total statement volume, it does not exhibit the

[3]Appendix A.4 documents that unusually long statements in the corpus are concentrated in countries where the policy decision is communicated through a broader analytical document rather than a concise press release, reflecting institutional differences in communication formats and document bundling.

longest historical coverage. This institutional variation reflects differences in transparency regimes, governance structures, and the evolution of communication strategies among monetary authorities.

Figure 3: Statement Volume and Coverage by Country

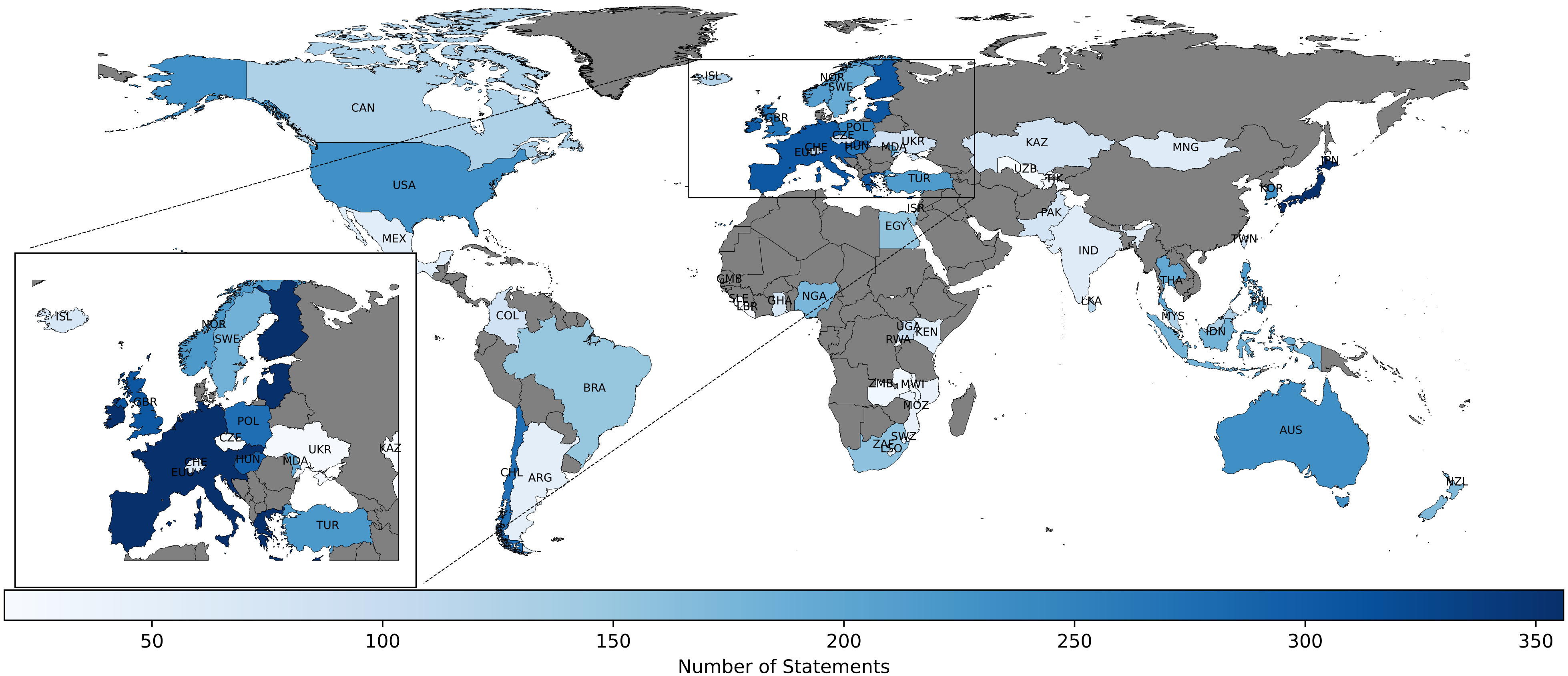


(a) Statements per Country **(b)** Coverage Period (Years)

Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

The map in Figure 4 illustrates the global reach of the MPSD, encompassing central banks from advanced economies, emerging markets, and developing countries across different continents. This geographical diversity is crucial for studying how monetary policy communication varies across different economic contexts and institutional settings, enabling more robust cross-country analyses than previously possible with existing datasets.

Figure 4: Number of Monetary Policy Statements by Country in the MPSD



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

Table 3: Central Bank Monetary Policy Statements: Coverage and Communication Frequency

Country	Region	Income Group	Number of Statements	Avg. Statements per Year	Modal Statements per Year	Statements in 2023	Database
Japan (JPN)	East Asia & Pacific	High income	356	13	8	8	Programmatic
Euro area (EUU)	Europe & Central Asia	High income	304	11	12	8	Programmatic
Chile (CHL)	Latin America & Caribbean	High income	274	10	12	8	Programmatic
United Kingdom (GBR)	Europe & Central Asia	High income	271	9	12	8	Programmatic
Hungary (HUN)	Europe & Central Asia	High income	264	11	12	12	Programmatic
Poland (POL)	Europe & Central Asia	High income	251	10	11	11	Programmatic
South Korea (KOR)	East Asia & Pacific	High income	233	9	12	8	Programmatic
Australia (AUS)	East Asia & Pacific	High income	231	6	11	11	Programmatic
United States (USA)	North America	High income	230	7	8	8	Programmatic
Turkey (TUR)	Europe & Central Asia	Upper middle income	215	11	12	12	Programmatic
Philippines (PHL)	East Asia & Pacific	Lower middle income	215	9	8	9	Programmatic
Norway (NOR)	Europe & Central Asia	High income	214	7	8	8	Programmatic
Moldova (MDA)	Europe & Central Asia	Upper middle income	200	8	12	9	Programmatic
Thailand (THA)	East Asia & Pacific	Upper middle income	194	7	8	6	Programmatic
Sweden (SWE)	Europe & Central Asia	High income	191	6	6	5	Programmatic
Indonesia (IDN)	East Asia & Pacific	Upper middle income	181	9	12	12	Programmatic
Nigeria (NGA)	Sub-Saharan Africa	Lower middle income	175	8	11	9	Programmatic
New Zealand (NZL)	East Asia & Pacific	High income	171	6	8	7	Programmatic
South Africa (ZAF)	Sub-Saharan Africa	Upper middle income	159	6	6	6	Programmatic
Egypt (EGY)	Middle East & North Africa	Lower middle income	155	7	8	8	Programmatic
Brazil (BRA)	Latin America & Caribbean	Upper middle income	151	7	8	8	Programmatic
Israel (ISR)	Middle East & North Africa	High income	130	5	8	8	Programmatic
Malaysia (MYS)	East Asia & Pacific	Upper middle income	129	4	6	6	Programmatic
Canada (CAN)	North America	High income	128	8	8	8	Programmatic
Sri Lanka (LKA)	South Asia	Lower middle income	122	9	8	8	Manual

Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD). Region and Income Group data are from the World Bank country classifications by income level for 2022–2023.

Notes: This table presents statistics on monetary policy statements for the 25 central banks with more than 120 statements in our database. Coverage period spans from January 1990 through December 2024. "Programmatic" indicates that data are collected via a Python script through an automated interface that updates regularly, while "Manual" indicates manually collected data. Modal statements per year represent the most frequently occurring number of statements issued annually. Income group classifications follow World Bank definitions. Table 3 reports this information for the illustrative subsample discussed in the text, while the complete classification for all countries in the database is provided in the accompanying *metadata.csv* file in the public repository.

Table 4: Central Bank Monetary Policy Statements: Complete Database

Country	Code	Number of Statements	First Statement	Last Statement	Coverage (Years)	Database Type	Document Type	English Language	Central Bank	Region
Argentina	ARG	52	2016-08-09	2024-11-01	9	Programmatic	MPC	Translated	Central Bank of Argentina	Latin America & Caribbean
Australia	AUS	231	1990-01-23	2024-12-10	35	Programmatic	MPD / IRD	Official	Reserve Bank of Australia	East Asia & Pacific
Brazil	BRA	151	2006-03-08	2024-12-11	19	Programmatic	CS	Translated	Central Bank of Brazil	Latin America & Caribbean
Canada	CAN	128	2009-01-20	2024-12-11	16	Programmatic	MPD	Official	Bank of Canada	North America
Chile	CHL	274	1998-01-08	2024-09-03	27	Programmatic	PR of RPM	Translated	Central Bank of Chile	Latin America & Caribbean
Colombia	COL	85	2015-08-21	2024-10-31	10	Programmatic	PR	Translated	Central Bank of Colombia	Latin America & Caribbean
Czechia	CZE	88	2014-02-06	2024-11-07	11	Programmatic	BD	Translated	Czech National Bank	Europe & Central Asia
Egypt	EGY	155	2005-06-02	2024-11-21	20	Programmatic	MPCD	Translated	Central Bank of Egypt	Middle East & North Africa
Eswatini	SWZ	59	2011-05-13	2024-11-22	14	Manual	MPS	Official	Central Bank of Eswatini	Sub-Saharan Africa
Euro area	EUU	304	1999-03-04	2024-12-12	26	Programmatic	IS / MPS	Official	European Central Bank	Europe & Central Asia
Gambia	GMB	37	2015-02-24	2024-11-26	10	Manual	MPC PR	Official	Central Bank of The Gambia	Sub-Saharan Africa
Ghana	GHA	71	2011-02-18	2024-11-29	14	Manual	MPC PR	Official	Bank of Ghana	Sub-Saharan Africa
Hungary	HUN	264	2002-12-16	2024-12-17	23	Programmatic	PR	Translated	Magyar Nemzeti Bank	Europe & Central Asia
Iceland	ISL	120	2009-01-29	2024-11-20	16	Programmatic	Statement of MPC	Official	Central Bank of Iceland	Europe & Central Asia
India	IND	59	2015-04-07	2024-12-06	10	Programmatic	MPS	Official	Reserve Bank of India	South Asia
Indonesia	IDN	181	2005-08-16	2024-11-20	20	Programmatic	PR	Translated	Bank Indonesia	East Asia & Pacific
Israel	ISR	130	2002-09-23	2024-11-25	23	Programmatic	IRA	Translated	Bank of Israel	Middle East & North Africa
Japan	JPN	356	1998-01-16	2024-12-19	27	Programmatic	Statement on MP	Translated	Bank of Japan	East Asia & Pacific
Kazakhstan	KAZ	84	2006-09-05	2024-11-29	19	Programmatic	PR	Translated	National Bank of Kazakhstan	Europe & Central Asia
Kenya	KEN	55	2015-08-05	2024-12-05	10	Manual	PR	Official	Central Bank of Kenya	Sub-Saharan Africa
South Korea	KOR	233	1999-12-02	2024-11-28	26	Programmatic	MPD	Translated	Bank of Korea	East Asia & Pacific
Lesotho	LSO	43	2017-04-04	2024-11-26	8	Manual	Statement of MPC	Official	Central Bank of Lesotho	Sub-Saharan Africa
Liberia	LBR	18	2019-11-20	2024-08-27	6	Manual	Communique	Official	Central Bank of Liberia	Sub-Saharan Africa
Malawi	MWI	33	2014-04-30	2024-11-04	11	Manual	Statement of MPC	Official	Reserve Bank of Malawi	Sub-Saharan Africa
Malaysia	MYS	129	1998-02-06	2024-11-06	26	Programmatic	MPS	Translated	Central Bank of Malaysia	East Asia & Pacific
Mexico	MEX	55	2018-05-17	2024-12-19	7	Programmatic	MPS	Translated	Bank of Mexico	Latin America & Caribbean
Moldova	MDA	200	2001-01-25	2024-12-05	24	Programmatic	MPD	Translated	National Bank of Moldova	Europe & Central Asia
Mongolia	MNG	59	2010-07-23	2024-12-13	15	Manual	MPS	Translated	Bank of Mongolia	East Asia & Pacific
Mozambique	MOZ	44	2017-10-26	2024-11-27	8	Manual	Communique	Translated	Bank of Mozambique	Sub-Saharan Africa
New Zealand	NZL	171	1999-03-17	2024-11-27	26	Programmatic	MPD	Official	Reserve Bank of New Zealand	East Asia & Pacific
Nigeria	NGA	175	2004-02-18	2024-11-26	21	Programmatic	MP Communique	Official	Central Bank of Nigeria	Sub-Saharan Africa
Norway	NOR	214	1996-11-05	2024-12-19	29	Programmatic	Rate decision	Translated	Central Bank of Norway	Europe & Central Asia
Pakistan	PAK	79	2005-11-27	2024-07-29	20	Programmatic	MPS	Official	State Bank of Pakistan	South Asia
Philippines	PHL	215	2002-01-17	2024-12-19	23	Programmatic	MPD	Official	Central Bank of the Philippines	East Asia & Pacific
Poland	POL	251	2001-11-28	2024-12-04	24	Programmatic	MPCo PR	Translated	Narodowy Bank Polski	Europe & Central Asia
Rwanda	RWA	21	2019-02-07	2024-11-20	6	Manual	PR	Official	National Bank of Rwanda	Sub-Saharan Africa
Sierra Leone	SLE	20	2015-12-14	2024-12-23	10	Manual	MPS	Official	Bank of Sierra Leone	Sub-Saharan Africa
South Africa	ZAF	159	1999-10-13	2024-11-21	26	Programmatic	Statement of MPC	Official	South African Reserve Bank	Sub-Saharan Africa
Sri Lanka	LKA	122	2012-05-11	2024-11-27	13	Manual	MP Review	Translated	Central Bank of Sri Lanka	South Asia
Sweden	SWE	191	1996-01-09	2024-12-19	29	Programmatic	MPD	Translated	Sveriges Riksbank	Europe & Central Asia
Switzerland	CHE	110	2000-01-20	2024-06-20	25	Programmatic	MPA	Translated	Swiss National Bank	Europe & Central Asia
Taiwan	TWN	98	2001-06-28	2024-09-19	24	Programmatic	MPD	Translated	Central Bank of the Republic of China (Taiwan)	East Asia & Pacific
Tajikistan	TJK	23	2017-01-31	2024-04-26	8	Manual	PR	Translated	National Bank of the Republic of Tajikistan	Europe & Central Asia
Thailand	THA	194	2000-05-23	2024-12-18	25	Programmatic	MPCD	Translated	Bank of Thailand	East Asia & Pacific
Turkey	TUR	215	2006-01-23	2024-08-20	19	Programmatic	PR on Interest Rates	Translated	Central Bank of the Republic of Turkey	Europe & Central Asia
Uganda	UGA	97	2011-07-01	2024-12-05	14	Manual	MPS	Official	Bank of Uganda	Sub-Saharan Africa
Ukraine	UKR	87	2014-04-15	2024-12-13	11	Programmatic	MPD	Translated	National Bank of Ukraine	Europe & Central Asia
United Kingdom	GBR	271	1997-06-06	2024-11-07	28	Programmatic	MPSm	Official	Bank of England	Europe & Central Asia
United States	USA	230	1994-02-04	2024-12-18	31	Programmatic	FOMC statement	Official	Board of Governors of the Federal Reserve System	North America
Uzbekistan	UZB	21	2020-01-18	2024-12-12	5	Manual	PR	Translated	Central Bank of the Republic of Uzbekistan	Europe & Central Asia
Zambia	ZMB	31	2016-11-16	2024-11-13	9	Manual	MPC Statement	Official	Bank of Zambia	Sub-Saharan Africa

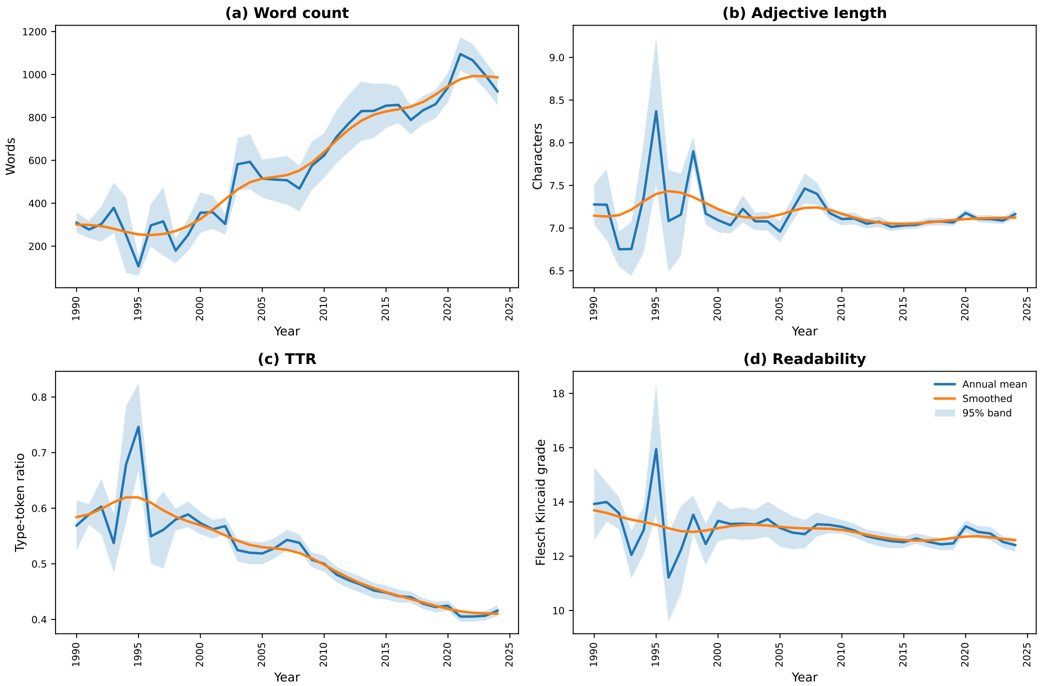
Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD). Region data are from the World Bank country classifications.

Notes: This table presents the complete coverage of the MPSD across all 51 central banks. Document Type abbreviations: MPC = Monetary Policy Committee; MPD = Monetary Policy Decision; MPS = Monetary Policy Statement; IS = Introductory Statement; PR = Press Release; CS = Copom Statement; BD = Board Decision; IRD = Interest Rate Decision; MPCD = Monetary Policy Committee Decision; IRA = Interest Rate Announcement; MPA = Monetary Policy Assessment; RPM = Rate-Setting Policy Meeting; MPCo = Monetary Policy Council; MP = Monetary Policy. Coverage years are calculated as the time between the first and last statements in our database.

3.3 Linguistic Patterns

We characterize the evolution of monetary policy statements using four complementary linguistic measures that capture distinct dimensions of central bank communication: scale, morphological complexity, lexical diversity, and readability (Figure 5).

Figure 5: Evolution of Statement Length and Linguistic Characteristics



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

Notes: Four document-level linguistic indicators computed for each statement: (i) word count (total tokens); (ii) average adjective length (mean character count of adjectival tokens via part-of-speech tagging); (iii) type-token ratio (unique tokens divided by total tokens); and (iv) Flesch-Kincaid Grade Level. Statement-level measures are aggregated annually using cross-statement means. Shaded bands represent 95% confidence intervals based on the standard error of the annual mean. Smoothed trends obtained via Gaussian filtering.

Panel (a) documents a pronounced upward trend in average word count, with statements approximately tripling in length between 1990 and 2024. The expansion accelerates markedly following the Global Financial Crisis, consistent with the adoption of unconventional monetary policies and forward guidance, which require more extensive explanation. This finding aligns with Masciandaro et al. (2024)'s survey documenting the increasing complexity of central bank communication frameworks. Panel (b) reports average adjective length—a measure of morphological complexity in evaluative language. This indicator remains stable over the sample period, fluctuating modestly around 7.5 characters. The stability suggests that while statements have expanded substantially, central banks have not systematically adopted more complex descriptive vocabulary;

lengthening occurs primarily through additional content rather than denser prose. Panel (c) displays the type-token ratio (TTR), a standard measure of lexical diversity. The secular decline in TTR—from approximately 0.60 in the early 1990s to 0.40 by 2024—indicates increasing lexical repetition as statements lengthen. This pattern is consistent with the institutionalization of standardized policy language and templated communication structures documented in Hansen and McMahon (2016). Panel (d) presents Flesch-Kincaid Grade Level scores, where lower values indicate greater accessibility. Despite the substantial increase in length, readability improves modestly over time, suggesting that central banks have partially offset greater informational content with simpler sentence structures and more accessible word choices. This evolution reflects deliberate efforts to reach broader audiences beyond financial market specialists. Collectively, these patterns reveal a fundamental transformation in central bank communication: statements have become longer and more standardized while maintaining—and even improving—accessibility. This shift toward detailed yet readable policy communication represents a purposeful adaptation to the expanded role of transparency in modern monetary policy frameworks.[4]

A closer examination of monetary policy statements reveals a typical structure including economic forecasts, domestic and foreign economic analysis, policy announcements regarding interest rates and forward guidance, and comments on the policy mix. These statements exhibit recurring sentence patterns, often involving specific parts-of-speech sequences, as detailed in Table 5.

Table 5 illustrates the systematic relationship between economic themes and linguistic structures in monetary policy statements. By identifying these recurring syntactic patterns, we develop more effective NLP tools tailored to central bank communications. For instance, the sequence "policy + AUX + ADJ" is frequently associated with the Monetary Stance, as seen in examples like "policy is restrictive" or "stance remains loose," while the structure "NOUN + ADP + inflation" typically characterizes Inflation Dynamics. These patterns, identified using spaCy's Universal POS tags combined with lexical anchors, provide a robust framework for disentangling the stylistic conventions of central bank communication from simple word frequency.

Figure 6 highlights the most frequent part-of-speech patterns observed in monetary policy statements, providing further evidence of the structured nature of central bank communications.

As shown in Figure 6, common sentence patterns include `DET ADJ NOUN` (e.g., "The strong dollar," "The current rate"), `DET NOUN ADP` (e.g., "The impact of," "The level of"), and `ADP DET NOUN` (e.g., "In the market," "On the economy"). The prevalence of these patterns suggests a high degree of formalization in monetary policy discourse, with central banks employing consistent linguistic structures to

[4]Appendix A.3 presents these linguistic indicators separately for statements originally drafted in English and those translated into English. The similarity of trends across both groups suggests that the main results are not driven by translation-related measurement artifacts.

Table 5: Linguistic Patterns in Monetary Policy Statements by Economic Theme

Economic Theme	**Part-of-Speech Pattern**	**Representative Example**
Inflation Forecasts	inflation + AUX + VERB	inflation **will rise** inflation **might persist**
Growth Projections	growth/GDP + AUX + VERB	growth **should recover** GDP **may stall**
Inflation Dynamics	NOUN + ADP + inflation	risk **of** inflation outlook **for** inflation
Rate Levels	rate + ADP + NUM	rate **at** 5.25% rate **above** 2.0
Forward Guidance	policy/rate + AUX + VERB	policy **will remain** rates **might stay**
Monetary Stance	policy/stance + AUX + ADJ	policy **is restrictive** stance **remains loose**
Fiscal Policy	NOUN + ADP + spending/deficit	impact **of** deficit boost **from** spending
External Conditions	NOUN + ADP + trade/export	demand **for** exports uncertainty **about** trade

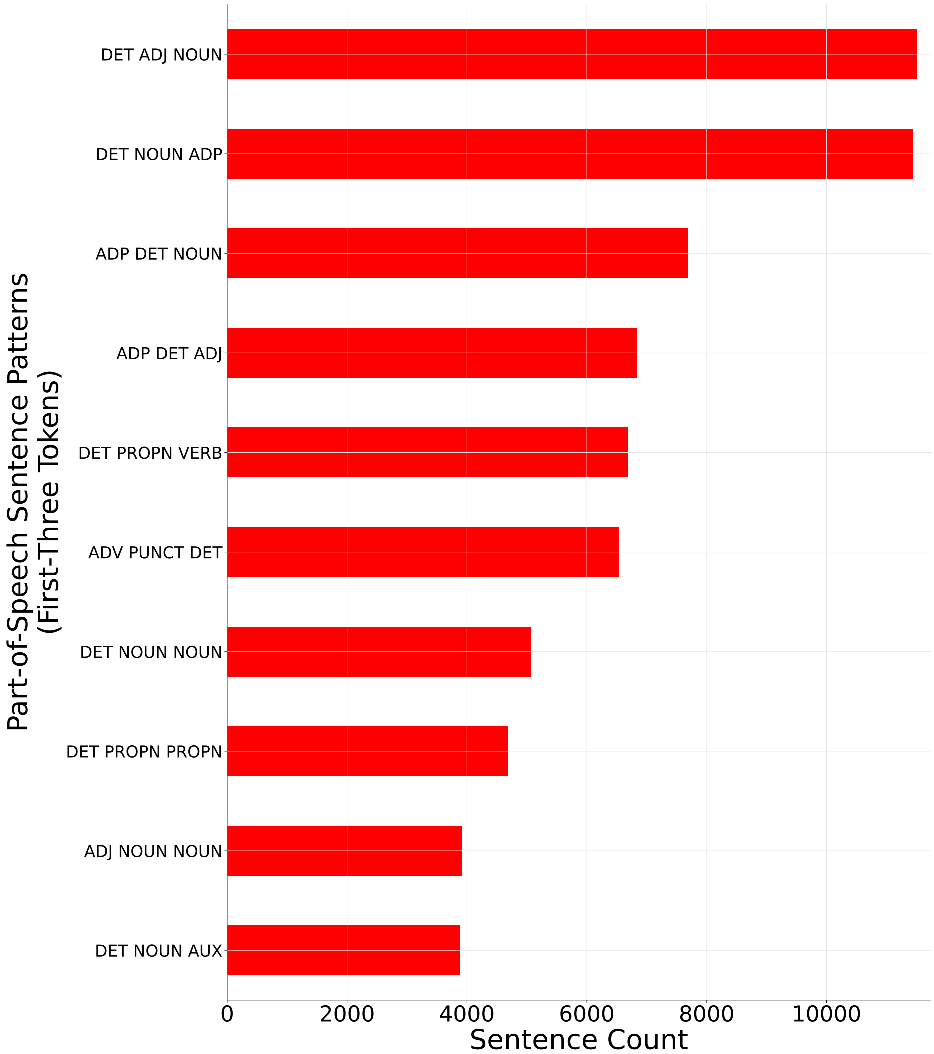
Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

Notes: This table presents the simplified recurring linguistic patterns identified in our corpus. Pattern identification employs a structural approach combining Universal POS tags (AUX = Auxiliary, ADP = Adposition, ADJ = Adjective, NUM = Number) with single lemma-based lexical anchors. This structure allows for the capture of complex syntactic relationships (e.g., "risk of inflation") rather than simple word adjacency.

convey their assessments and decisions. Additionally, the `DET PROPN VERB` pattern is particularly associated with central bank commentary and actions, as illustrated in Appendix Figure 20.

Figure 7 presents a word cloud visualization of terms appearing in the `NOUN + ADP + inflation` pattern as shown in Table 5. This structure effectively filters for the context of inflation rather than mere mentions of the word. Dominant trigrams such as "risks to inflation," "outlook for inflation," and "decline of inflation" highlight that this syntactic pattern is primarily used to communicate uncertainty and future trajectory. The high frequency of terms like "rise," "pressures," and

Figure 6: Most Common Patterns by Part-of-Speech (POS)



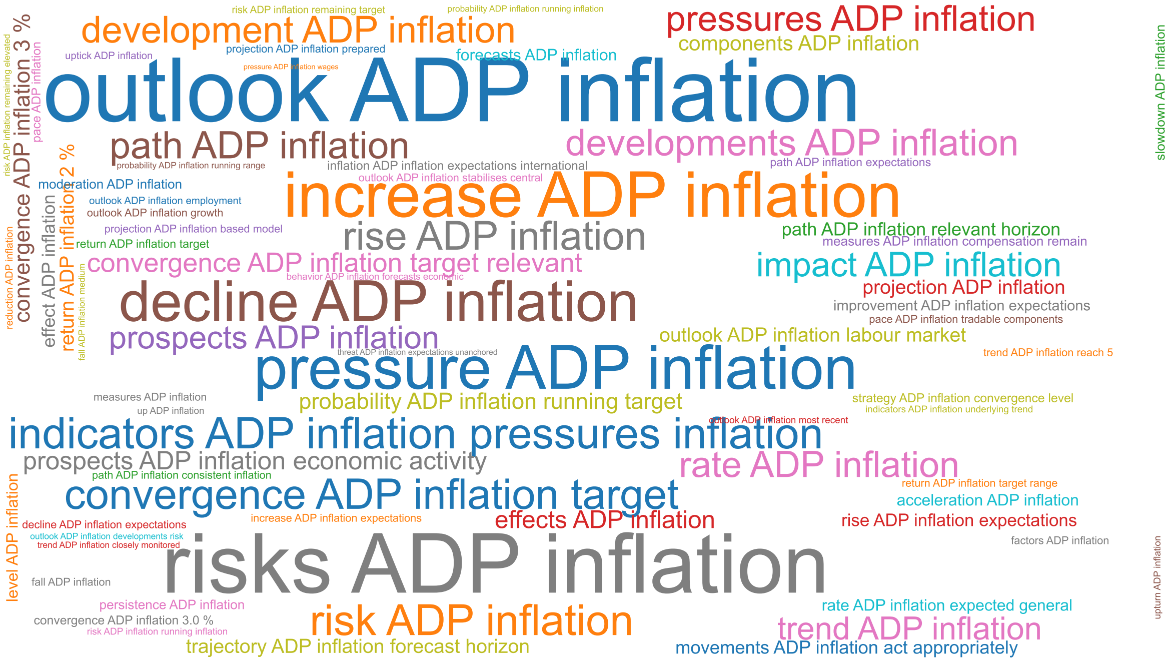
Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

"convergence" confirms that this pattern acts as a linguistic marker for the central bank's assessment of price dynamics.

3.4 Corpus Characteristics and Word Frequency

Figure 8 presents a word cloud visualization of the most frequent terms in our corpus of central bank communications. The prominence of terms such as "inflation," "economic," "growth," "policy," and "financial" reflects the core concerns of monetary policymakers. This visualization not only illustrates the thematic focus of central bank communications but also demonstrates the richness and domain-specificity of our corpus, which is essential for training specialized language models that can accurately capture the nuances of monetary policy discourse.

Figure 7: Most Common Trigrams with the Part-of-Speech Pattern `NOUN + ADP + inflation`



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

Notes: The inflation dynamics (`NOUN + ADP + inflation`) pattern highlights contextual usage such as "risks to inflation" and "outlook for inflation." Matches: 7,899

3.5 Inflation Mentions Over Time

We analyze the evolution of key economic terms in monetary policy communications. Figure 9 tracks references to inflation across central banks over time, providing a measure of the prominence of price stability in policy discourse. To account for mechanical variation arising from changes in communication frequency and document length, we report two complementary normalized measures. Panel (a) of Figure 9 shows inflation mentions normalized by the number of statements issued per year, capturing the average intensity of inflation-related discussion per policy communication. Panel (b) normalizes inflation mentions by total word count, reporting mentions per 1,000 words and thus measuring the relative salience of inflation within the textual content of policy statements.

These normalizations ensure that observed changes in inflation-related communication are not mechanically driven by an increased number of policy meetings or longer statements, but instead reflect shifts in the emphasis placed on inflation by central banks. The resulting series reveal substantial comovement in inflation discourse across countries, with particularly pronounced increases during episodes of elevated global inflationary pressure. At the same time, persistent cross-country differences remain, reflecting heterogeneity in mandates,

Figure 8: Standardized and Diverse Corpus for Training/Testing A.I.

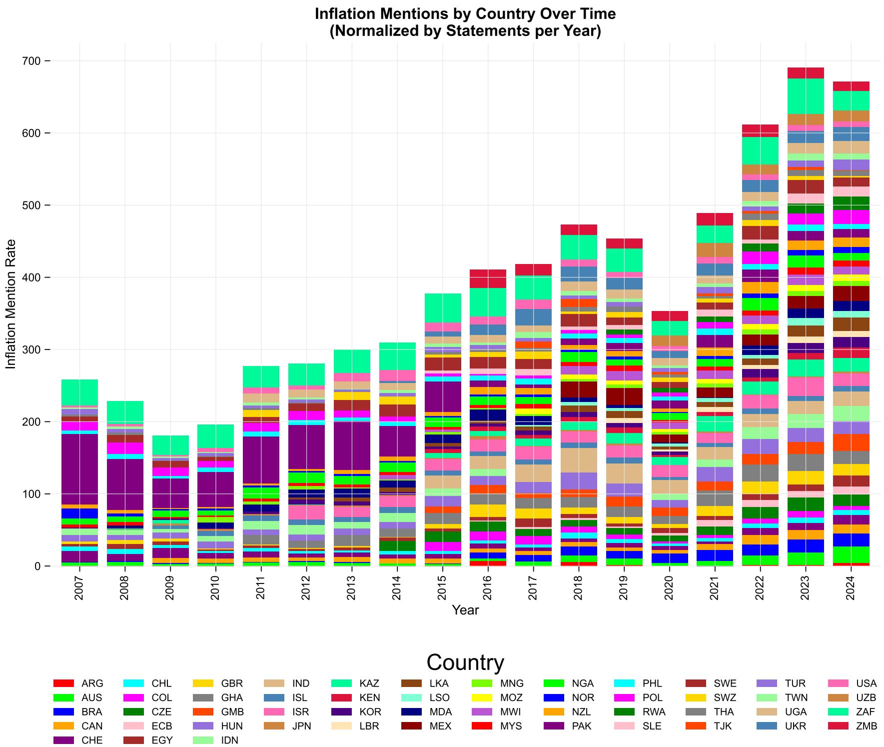
Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

Notes: 51 central banks and 200,456 sentences.

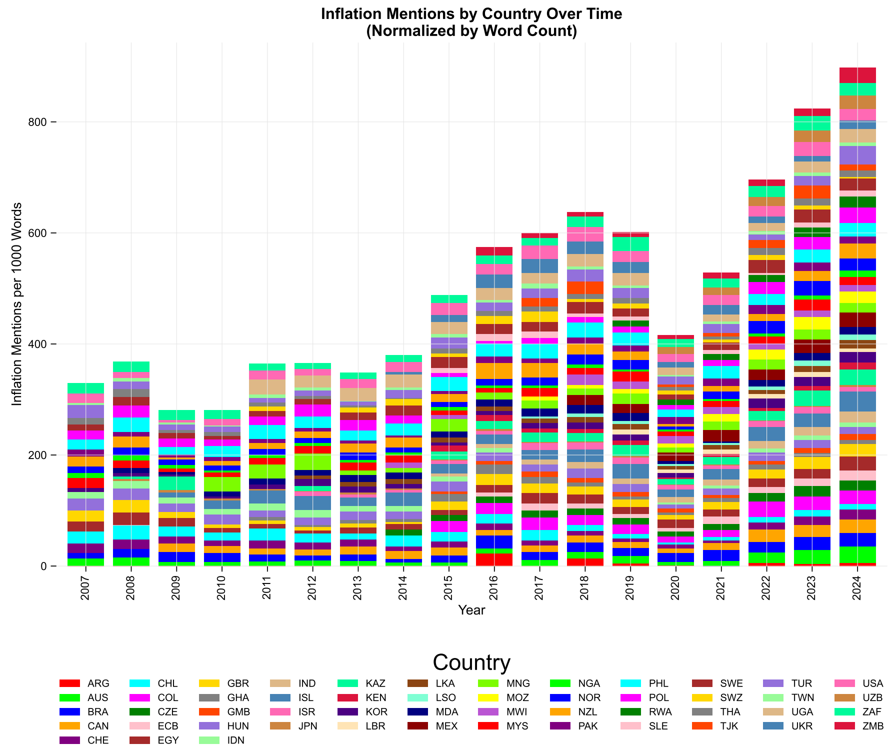
inflation dynamics, and policy frameworks. Overall, the normalized measures provide a clearer interpretation of how central banks' communication priorities evolve in response to changing macroeconomic conditions, while facilitating meaningful comparison across institutions and over time.

This analysis of monetary policy statements contributes to a deeper understanding of how central banks communicate their policy stances and perspectives on economic developments. By dissecting the structure, content, and evolution of these statements, we gain valuable insights into the strategies and objectives of central banks.

Figure 9: Inflation Mentions by Country



(a) Normalized by statements per year



(b) Normalized by word count

Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

4 A Generative AI Application to Central Bank Communications

This section illustrates the capabilities of generative AI, specifically LLMs, to analyze economic data and address pertinent challenges in economic research. Our application focuses on examining the GFC, a phenomenon characterized by the synchronized movement of capital, credit, and asset prices across countries (Rey, 2015; Miranda-Agrippino and Rey, 2020, 2022).

The applications presented here are illustrative and are intended to demonstrate how the database and infrastructure can be used, rather than to establish definitive causal claims.

4.1 Conceptual Framework

Central bank communications influence global financial conditions through multiple channels. First, statements from major central banks—particularly the Federal Reserve and ECB—serve as public signals about the state of the global economy, reducing information asymmetries among market participants worldwide (Morris and Shin, 2002). Second, policy communications coordinate expectations across borders, affecting asset prices and capital flows even in countries not directly targeted by the communication (Woodford, 2005). Third, the content and tone of central bank statements provide forward guidance that shapes expectations about future monetary policy paths globally (Blinder et al., 2008). These mechanisms suggest that systematically analyzing central bank communications can reveal how policymakers perceive and respond to global economic dynamics—precisely the objective our infrastructure enables.

4.2 Addressing Challenges in AI Applications to Economics

While LLMs offer immense analytical potential, they present significant methodological challenges for economic research. We identify and address two primary concerns:

1. **Reproducibility**: The lack of transparency surrounding the inputs and internal mechanisms of many LLMs hinders the replication of research findings, a cornerstone of the scientific method. To address this, our study prioritizes open-source code, providing access to LLM code inputs, prompts used for model querying, a clean code base, and user-friendly Jupyter notebooks. This approach ensures complete transparency and facilitates the reproduction of the research process, promoting rigorous scientific inquiry.

2. **Computational Efficiency**: High computational costs and energy consumption associated with LLMs present practical barriers to their

application in economic research. To mitigate this, we implement text-filtering pipelines that leverage keyword and part-of-speech analysis to refine the input data for the LLMs, effectively reducing the computational burden and enhancing efficiency.

While the internal mechanisms of LLMs are not fully transparent, reproducibility in empirical economics relies on the ability to replicate results using the same data, code, and documented procedures. We therefore ensure transparency at every stage under the researcher's control, including data collection, preprocessing, prompt design, model selection, and analytical workflows. This approach is consistent with current best practices in text-as-data research using modern NLP tools (Gentzkow et al., 2019). Importantly, scientific reproducibility does not require replicating the original data collection process; rather, it requires that published results can be exactly reproduced from the shared corpus and code. All figures, tables, and empirical findings in this paper satisfy this criterion.

Table 6 illustrates the output of our LLM-based analysis of global financial conditions over time. The visualization shows how central banks' discussion of global economic drivers has evolved through different economic cycles. This approach demonstrates how LLMs can systematically extract and analyze economic narratives from large corpora of policy communications, offering insights into how policymakers conceptualize and respond to global economic dynamics. The temporal patterns revealed in this analysis align with major economic events and shifts in the GFC, highlighting the value of applying advanced language models to economic research questions.

Table 6: LLM-Generated Analysis of Global Financial Conditions (2007–2013)

Year and Driver	Origin Country	Summary
2007 Global economic conditions, subprime mortgage crisis, global liquidity, U.S. economic slowdown/housing market, commodity prices.	U.S., China	Several central banks identify global economic conditions as a key driver, with concerns arising from the U.S. economic slowdown/housing market and the subprime mortgage crisis. Concerns regarding global liquidity and commodity prices (oil and food) are also noted. The U.S. Federal Reserve is implicitly referenced through discussions of the U.S. economy and monetary policy, but no central bank explicitly identifies the Federal Reserve as the main driver.
2008 U.S. subprime mortgage crisis, Global financial turmoil, Slowdown in the U.S. economy, Global economy slowdown.	United States	The main driver of global economic and financial conditions is the United States, specifically originating from the U.S. subprime mortgage crisis.
2009 Global Financial Crisis / Global Economic Slowdown / Global Recession.	United States (due to U.S. housing bust)	The main driver of global economic and financial conditions during this period was identified as the Global Financial Crisis, originating from the U.S. housing bust. Many central banks acknowledged the significant impact of the crisis on their domestic economies. No central banks directly identified the Federal Reserve as the driver, but the origin of the crisis is cited as the U.S. housing market.
2010 Global Economic Recovery.	Not explicitly one country (mentions of U.S., China, Asia)	The primary driver of global economic and financial conditions is identified as the global economic recovery, influenced by emerging market economies (especially China) and, to a lesser extent, advanced economies like the U.S. and Japan. Multiple texts mention fiscal stimulus measures as supporting the recovery.
2011 Sovereign debt crisis in Europe and global economic slowdown.	Euro Area	The main driver of global economic and financial conditions is the sovereign debt crisis in Europe, which is causing a global economic slowdown. Several central banks identify the European debt crisis as a significant risk. The Federal Reserve's Quantitative Easing (QE) is mentioned by Bank Indonesia as an event to monitor, but it is not portrayed as the primary driver of the global economy.
2012 Sovereign Debt Crisis in Europe / Weak Global Growth.	Euro Area / Multiple Countries	The main driver of global economic and financial conditions appears to be the sovereign debt crisis in Europe and the overall weak global growth. Many central banks across the globe cite the Eurozone crisis as a major source of risk and uncertainty. The central banks do not directly identify the Federal Reserve as the main driver of global economic and financial conditions, although some mention monetary policies of major central banks.
2013 Uncertainty regarding U.S. monetary policy (specifically tapering), fiscal policy, and debt ceiling issues.	United States (primarily), Euro Area and China	Several central banks identify the U.S. Federal Reserve's potential tapering of quantitative easing as a major driver of global financial market volatility and a downside risk to growth. The U.S. fiscal situation and debt ceiling debates also contribute to uncertainty. Concerns about the Euro Area's economic challenges and China's growth slowdown are also frequently mentioned.

Source: Structured summaries derived from the Monetary Policy Statement Database. We filtered 2,921 unique sentences containing the keywords "global" and "spillover" by year using spaCy (see Section 4.2) and generated summaries via Google's Gemini (gemini-2.0-flash) using the prompt in Appendix A.2. All input texts (sentences) are stored in JSON for replication and will be made available along with database metadata.

4.3 LLM Data Pipeline Infrastructure

Our methodological framework centers on a scalable LLM-based data pipeline architecture that integrates diverse language models—both proprietary (Claude, Gemini, ChatGPT) and open-source (hosted on Hugging Face)—within a unified computational environment. The pipeline's modular design permits seamless substitution of text corpora and models, enabling comparative analysis across central bank communications and addressing potential concerns regarding model-specific biases. This flexibility strengthens the external validity of our findings and ensures the analytical framework can seamlessly integrate emerging language models and additional data sources.

The architecture is designed for scalability along two dimensions:

- **LLM Agnosticism**: The system is compatible with any model (e.g., ChatGPT, Gemini, Claude) and task (e.g., sentiment analysis, QA)
- **Text/Prompt Flexibility**: The pipeline can be dynamically repurposed to evaluate alternative texts and prompts

4.4 Methodological Approach

To illustrate the application of LLMs to central bank communications, we explore the GFC through two key analytical approaches:

4.4.1 Question Answering

In the QA task, we prompt the LLM to provide summaries of the drivers of the global economy, with particular attention to the role of the Federal Reserve. The core directive of the prompt (the full prompt is available in Appendix A.2) is summarized as follows:

> "I want you to identify the main driver of global economic and financial conditions and the origin country driving these conditions. In the summary, please note if any central banks identify the Federal Reserve as the driver of global economic or financial conditions."

Table 7 presents a comparative analysis of how different LLMs interpret the *same* central bank text. This side-by-side comparison reveals significant variations in the summaries generated by different models, even when processing identical inputs and the same prompt. Notably, Claude identifies the U.S. Federal Reserve as the specific driver, whereas all other models—including Gemini (see Table 6)—attribute the influence to the U.S. economy more generally. These differences underscore the importance of transparency in LLM-based research and the need for careful model selection and evaluation when applying these tools to economic analysis. The figure illustrates why our pipeline's ability to integrate multiple models is valuable—it allows researchers to compare outputs across different LLMs and assess the robustness of their findings.

Table 7: Comparative Analysis of LLM Interpretations of 2008 Global Economic Conditions

Model	Identified Mechanism	Origin Country	Summary Assessment
Claude	The Federal Reserve	United States	Based on central bank statements, the main driver is the Federal Reserve. Several central banks (e.g., Canada, New Zealand, Australia) identified the Fed's actions as a key factor impacting global growth. The global slowdown stemming from the U.S. housing bust was a recurring theme.
Xiaomi	Global financial market turmoil and economic slowdown	United States	The primary driver is global financial market turmoil stemming from the U.S. subprime mortgage crisis. The Federal Reserve is acknowledged by multiple central banks (e.g., PK, ZA, NG, GB, ECB, JP) as a key driver given its role in global stability.
Mistral	Slowdown in the U.S. economy / subprime mortgage crisis	United States	Multiple central banks cite the US economic slowdown—particularly the subprime crisis—as the primary driver. Notably, no central bank explicitly names the Federal Reserve as the driver in this interpretation.
Arcee	U.S. subprime mortgage crisis and broader U.S. economic weakness	United States	Multiple central banks explicitly cite the U.S. subprime crisis and economic slowdown as primary drivers. Several banks note that the Federal Reserve's policy actions are relevant to addressing these conditions.

Source: Structured summaries derived from the Monetary Policy Statement Database. We filtered 2,921 unique sentences containing the keywords "global" and "spillover" by year using spaCy (see Section 4.2) and generated summaries via Anthropic's Claude, mistralai/devstral-2512, arcee-ai/trinity-mini, xiaomi/mimo-v2-flash. All input texts (sentences) are stored in JSON for replication and will be made available along with database metadata.

4.4.2 Aspect-Based Sentiment Analysis

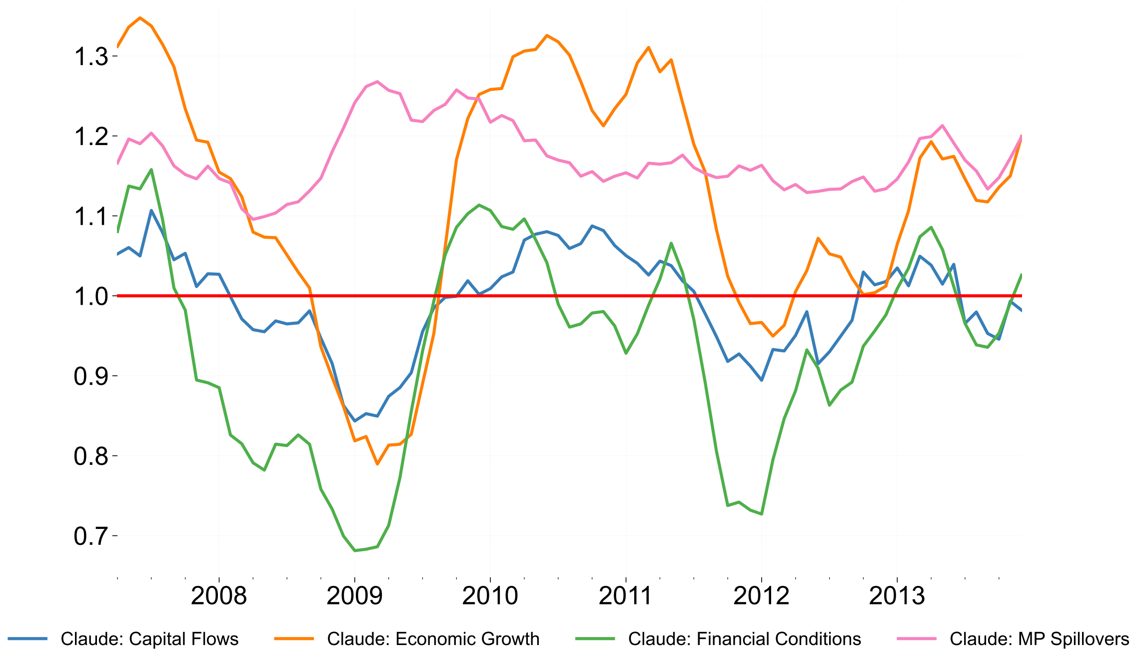
For the ABS task, we instruct the LLM to extract sentiment toward specific themes from central bank communications. The core directive of the prompt (the full prompt is available in Appendix A.1) is summarized as follows:

> "Analyze the provided central bank text to identify discussions regarding the impact of global financial conditions on domestic economies, focusing on themes such as economic growth, financial conditions, capital flows, and monetary policy spillovers. For each explicitly discussed theme, assign a continuous sentiment score on a

scale from 0.0 (negative) to 2.0 (positive), where 1.0 represents a neutral view."

This allows us to track sentiment toward specific economic factors across time and central banks, providing insights into how financial conditions and policy spillovers are perceived globally.

Figure 10: LLM-Generated ABS over Time



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD) and using LLM tools.

Notes: Sentiment scores are generated using a version of Anthropic's Claude LLM (claude-3-haiku-20240307). The series capture sentiment related to capital flows, economic growth, financial conditions, and monetary policy spillovers (MP Spillovers), averaged across all countries.

Figure 10 illustrates the evolution of sentiment across different economic aspects over time, as extracted by our LLM-based analysis. The figure demonstrates distinct patterns in how central banks express sentiment toward different themes, with notable divergences during periods of economic stress. For instance, the sentiment toward economic growth and financial conditions shows greater volatility compared to monetary policy spillovers and, to a lesser extent, capital flows, particularly during and after the Global Financial Crisis. These temporal patterns offer valuable insights into the shifting priorities and concerns of monetary policymakers across different economic cycles.

4.5 Validation of LLM-Based Sentiment

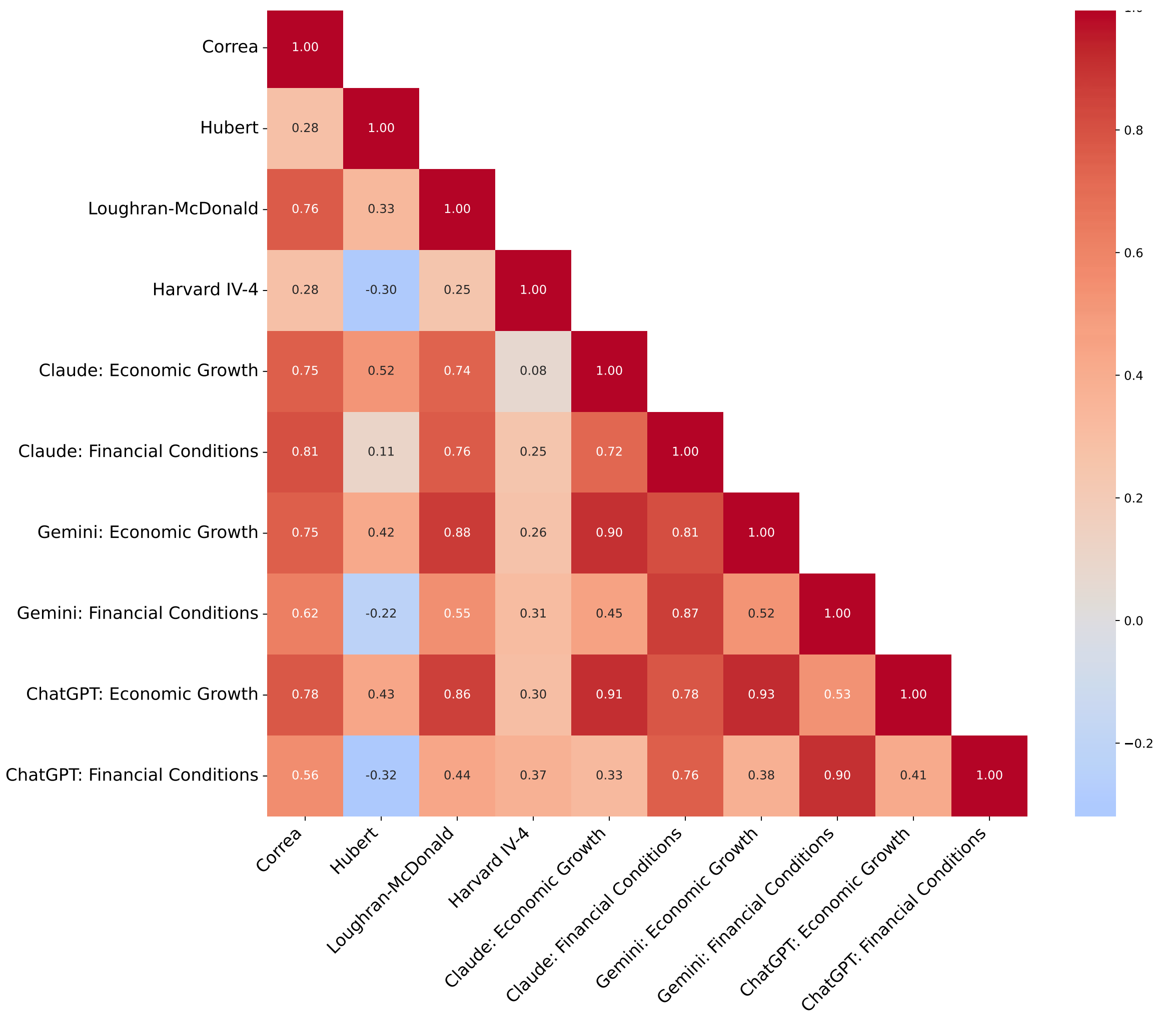
To assess the relationship between traditional lexical sentiment measures and LLM-derived sentiment indices, Figure 11 reports the correlation matrix across all sentiment metrics used in the analysis (see the extended matrix in Appendix Figure 18). Several clear patterns emerge. First, established finance-oriented lexical measures exhibit moderate to strong similarity among themselves. The Correa and Loughran-McDonald indices are strongly correlated (0.76), indicating that both capture a similar notion of financial tone in monetary policy statements. The Hubert index shows weaker alignment with these measures (correlations of 0.28 with Correa and 0.33 with Loughran-McDonald), while the general-purpose Harvard IV-4 dictionary displays substantially lower and less stable correlations, including a negative association with the Hubert index (-0.30). This pattern underscores the importance of domain-specific sentiment tools when analyzing central bank communication.

Second, LLM-derived sentiment measures show consistently strong correlations with finance-specific lexical indices, particularly for the Economic Growth dimension. Across models, Economic Growth sentiment correlates between 0.74 and 0.88 with the Loughran-McDonald index and between 0.75 and 0.78 with the Correa index. In contrast, correlations with the Harvard IV-4 dictionary remain markedly lower, typically below 0.26 (with even 0.08). This divergence suggests that LLMs align more closely with financial-domain sentiment constructs rather than with generic polarity-based lexicons.

Third, correlations differ systematically across sentiment dimensions. Financial Conditions sentiment exhibits weaker and more heterogeneous correlations with lexical measures than Economic Growth sentiment, with values ranging from 0.44 to 0.76 against Loughran-McDonald and from 0.56 to 0.81 against Correa. At the same time, Financial Conditions sentiment shows very strong agreement across LLMs themselves, with correlations exceeding 0.75 in most cases. This pattern indicates that while LLMs consistently identify a shared financial-conditions signal, that signal is only partially captured by traditional dictionary-based approaches.

Finally, the high cross-model correlations among LLMs—reaching 0.90–0.93 for Economic Growth sentiment—point to a robust consensus across different architectures and training paradigms. Rather than producing idiosyncratic sentiment measures, LLMs converge on similar interpretations of economic tone, reinforcing their reliability as measurement tools. Taken together, these results suggest that LLM-based sentiment measures are neither orthogonal to nor simple replications of traditional lexical indices. Instead, they complement established finance-specific dictionaries by capturing structured, dimension-specific sentiment—particularly for growth and financial conditions—while diverging from broad, non-domain-specific sentiment constructs. This finding aligns with the view that monetary policy communication conveys multiple informational

Figure 11: Correlation Matrix of Lexical and LLM-Based Sentiment Measures



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD) and using LLM tools and Loughran-McDonald, Correa, Hubert, and Harvard IV-4 dictionaries.

Notes: Sentiment correlation by method. Sentiment scores are generated using Loughran-McDonald, Correa, Hubert, and Harvard IV-4 dictionaries, and a version of Anthropic's Claude LLM (claude-3-haiku-20240307), OpenAI's ChatGPT (gpt-4o-mini) and Google's Gemini (gemini-2.0-flash). Correlations are computed based on the cross-sectional mean of the sentiment scores, aggregated across all sample countries.

dimensions and that modern LLMs are well-suited to disentangling these dimensions in a way that remains consistent with established financial text-analysis benchmarks.

These results contribute to the literature on text-as-data benchmarks. While Gentzkow et al. (2019) caution that different measures often target different estimands, our findings demonstrate reassuring convergence when the domain is strictly defined. The results support the premise of Loughran and McDonald (2011) regarding the necessity of domain-specific sentiment lexicons: the LLMs naturally emulate the financial-specific weighting of the Loughran–McDonald dictionary rather than the generic Harvard IV-4. Consequently, while Hansen and McMahon (2016) note that policy statements contain multiple informational dimensions, our analysis suggests that modern LLMs are highly effective at identifying the specific dimensions of financial risk and economic growth that traditional financial dictionaries were designed to capture.

We identify three mechanisms through which LLMs outperform dictionary methods in this setting. First, LLMs capture contextual negation and qualification that bag-of-words approaches miss. When a central bank states that "inflation pressures have not materialized as expected," dictionary methods count "inflation" and "pressures" as negative, while LLMs correctly interpret the overall sentiment as positive (declining concern). Second, LLMs handle domain-specific terminology more flexibly. Central bank communications frequently use technical phrases ("quantitative easing," "forward guidance," "macro-prudential") whose sentiment depends heavily on context; LLMs infer meaning from surrounding text while dictionaries either ignore such terms or assign fixed scores. Third, LLMs can extract ABS—distinguishing concern about inflation from concern about growth—whereas most dictionary approaches produce a single aggregate score. This granularity is essential for understanding how central banks balance competing objectives in their communications. Our correlation analysis confirms these advantages: LLM-derived sentiment for specific dimensions (growth, financial conditions) correlates more strongly with the GFC than aggregate dictionary-based measures, suggesting that aspect-specific extraction captures economically meaningful variation that aggregate approaches obscure.

4.5.1 Validation Framework

Given the absence of labeled ground truth for ABS in central bank communication, we validate LLM-derived measures using two complementary strategies. First, we benchmark LLM-based sentiment indices against established dictionary-based sentiment measures commonly used in the economics literature. We interpret agreement as evidence of plausibility and divergence as reflecting differences in the underlying estimands rather than measurement error per se. Consistent with this framework, we find that our LLM-derived sentiment tracks finance-specific lexical measures closely, validating the models' ability to capture relevant economic tone. Figure 11 summarizes cross-method correlations between lexical sentiment measures and LLM-derived aspect sentiments, while Figure 12 compares their time-series behavior around major global episodes. Second, we assess the robustness of LLM-based sentiment by comparing outputs across multiple LLMs. Specifically, in addition to our baseline results based on Claude, we replicate the ABS extraction using alternative LLMs (Gemini and ChatGPT). Figure 17 reports the resulting sentiment series for key economic dimensions, including economic growth and financial conditions. Despite differences in scale and volatility across models, the series exhibit strong comovement and similar temporal dynamics, particularly around major macroeconomic events. This cross-model consistency provides an additional validation layer, indicating that the extracted sentiment patterns are not driven by idiosyncratic behavior of a single LLM architecture. Finally, to ground these statistical aggregates in the underlying text, we conducted a manual review of a random sample of statements using our validation tool. While this qualitative inspection does not constitute a formal scientific audit, we observed substantial agreement between LLM outputs and human interpretation of the underlying text. An illustrative example of this validation exercise is provided in the Appendix (Figure 21). Taken together, benchmarking against transparent lexical measures, comparing across independent LLMs, and spot-checking against human judgment enhances confidence in the stability and interpretability of the proposed sentiment indicators, while preserving their ability to capture contextual and aspect-specific information that is difficult to obtain using dictionary-based approaches alone.

We acknowledge three limitations inherent to LLM-based sentiment extraction. First, prompt sensitivity may affect results; we mitigate this concern by documenting all prompts in the repository and comparing outputs across different model architectures. Second, model updates may affect reproducibility over time; we therefore report exact model versions (e.g., claude-3-haiku-20240307, gemini-2.0-flash, gpt-5-nano) and archive all outputs. Third, LLMs may exhibit biases inherited from training data; our cross-model comparison suggests that such biases do not dominate, as models trained on different corpora produce highly correlated sentiment series. These limitations are common to all LLM-based text analysis and do not invalidate our findings; rather, they underscore the importance of the validation framework we employ.

4.6 Empirical Implementation

Our implementation employs a multi-stage filtering process to extract relevant narrative elements from the corpus:

1. Search using keywords (global, spillover, flows): yielding 6,135 sentences
2. Truncate for specific time periods (e.g., 2019 data): reducing to 1,228 sentences

This approach balances computational efficiency with analytical rigor, addressing the computational challenges of LLM analysis of large text corpora while ensuring systematic sampling for robust results.

4.7 Results

Our analysis using LLMs produces several important insights into the GFC:

4.7.1 Origins of Global Financial Conditions

Our QA analysis, shown in Table 6, indicates that central banks often cite the U.S. economy as a driver of global economic conditions, and in some instances, specifically mention the U.S. Federal Reserve as a factor influencing global economic and financial conditions, particularly during periods of high uncertainty such as 2007 and 2009. However, this is not a consistent trend, and in many years, broader economic factors in the U.S. and other regions, such as the Eurozone sovereign debt crisis and emerging market economies, are perceived as more impactful. This nuanced understanding of global economic influences, where the relative importance of different factors, including central bank actions, can vary over time, contrasts with some academic literature that tends to emphasize the direct and dominant influence of Fed policy on global financial conditions, suggesting a potential gap between academic models and policymakers' conceptualization of global economic linkages (Kalemli-Özcan, 2019).

4.7.2 Sentiment Analysis Results

The ABS analysis shows that central banks expressed greater concern about growth and financial conditions than monetary policy and capital flows in the post-2008 period, likely reflecting the lasting impact of the Global Financial Crisis and increased global interconnectedness.

Our Granger causality analysis follows the formal specification:

$$\mathrm{GFC}_t = \alpha + \sum_{k=1}^{p} \beta_k \mathrm{GFC}_{t-k} + \sum_{k=1}^{p} \gamma_k \mathrm{Sentiment}_{t-k} + \varepsilon_t \tag{1}$$

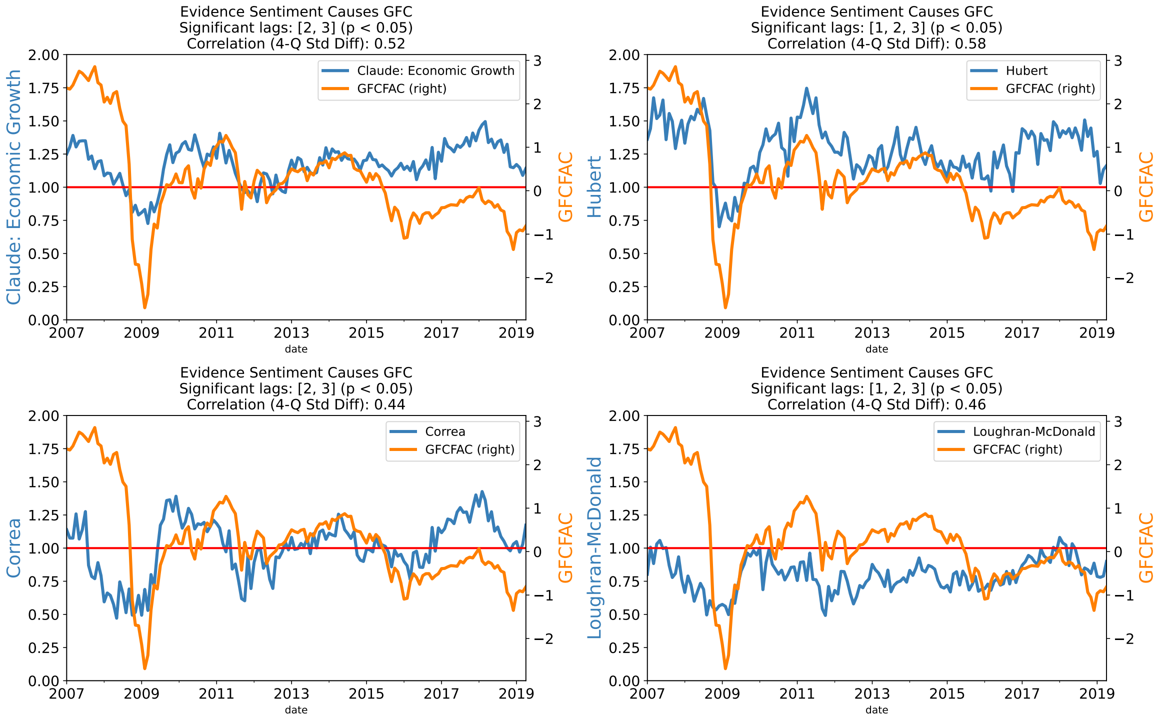
$$\mathrm{Sentiment}_t = \alpha' + \sum_{k=1}^{p} \beta'_k \mathrm{Sentiment}_{t-k} + \sum_{k=1}^{p} \gamma'_k \mathrm{GFC}_{t-k} + \varepsilon'_t \tag{2}$$

where GFC_t represents the GFC factor at time t, $\mathrm{Sentiment}_t$ represents our measure of central bank communication sentiment, and p represents the lag order. The null hypothesis that sentiment does not Granger cause GFC corresponds to $\gamma_k = 0$ for all $k \in \{1, 2, ..., p\}$. Similarly, the null hypothesis that GFC does not Granger cause sentiment corresponds to $\gamma'_k = 0$ for all $k \in \{1, 2, ..., p\}$.

We conduct these tests using lag orders $p \in \{1, 2, 3\}$ and report the resulting p-values in Figure 13. The results consistently reject the null hypothesis that sentiment does not Granger cause GFC, while failing to reject the null hypothesis that GFC does not Granger cause sentiment. This asymmetric causality pattern provides robust evidence that central bank communication sentiment influences global financial conditions rather than merely responding to them.

We emphasize that Granger causality is a statistical concept capturing predictive precedence rather than structural causation in the economic sense (Granger, 1969). The asymmetric pattern we document—where sentiment predicts GFC but not vice versa—is consistent with communication serving as a transmission channel, though we cannot rule out omitted common factors or reverse causality operating at frequencies not captured by our quarterly specification. Nevertheless, the finding that communication sentiment leads the GFC aligns with theoretical models in which central bank signals coordinate expectations and thereby influence asset prices (Morris and Shin, 2002; Woodford, 2005).

Figure 12: Comparison of LLM-Generated Sentiment Indices with Lexical Approaches



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD), dictionaries and LLM-based (Anthropic's Claude LLM claude-3-haiku-20240307) sentiment measures, and GFC factor (GFCFAC).

Notes: Global Sentiment vs. GFC. GFCFAC denotes the GFC factor. Blue lines show standardized global sentiment indices; orange lines show GFCFAC (right axis). Reported correlations refer to four-quarter standardized differences.

Figure 12 compares global sentiment measures derived from LLM-based and traditional lexical approaches with the GFC factor (GFCFAC). Contrary to the hypothesis that ABS measures might decouple from the aggregate cycle, the figure highlights a strong structural alignment between the methods. The LLM-based economic growth sentiment displays a robust correlation (0.52) with the GFC, effectively outperforming traditional finance-specific dictionaries such as Correa (0.44) and Loughran-McDonald (0.46), and performing comparably to the Hubert index (0.58). All measures exhibit significant predictive power at lags 2 and 3, confirming that the LLM-derived signal captures the same fundamental turning points as established benchmarks.

Taken together with Figures 10 and 11, these patterns indicate that LLM-based sentiment is highly consistent with the signal captured by domain-specific lexical dictionaries. Rather than diverging, the LLM-derived indices appear to refine the signal found in traditional measures. The fact that the LLM-based growth sentiment tracks the GFC more closely than the Loughran–McDonald and Correa indices suggests that the model's contextual awareness allows it to filter out noise better than bag-of-words approaches. Consequently, the observed convergence between these methods serves as a robust cross-validation, demonstrating that LLMs can reliably recover key macroeconomic dynamics while offering the additional advantage of aspect-specific granularity.

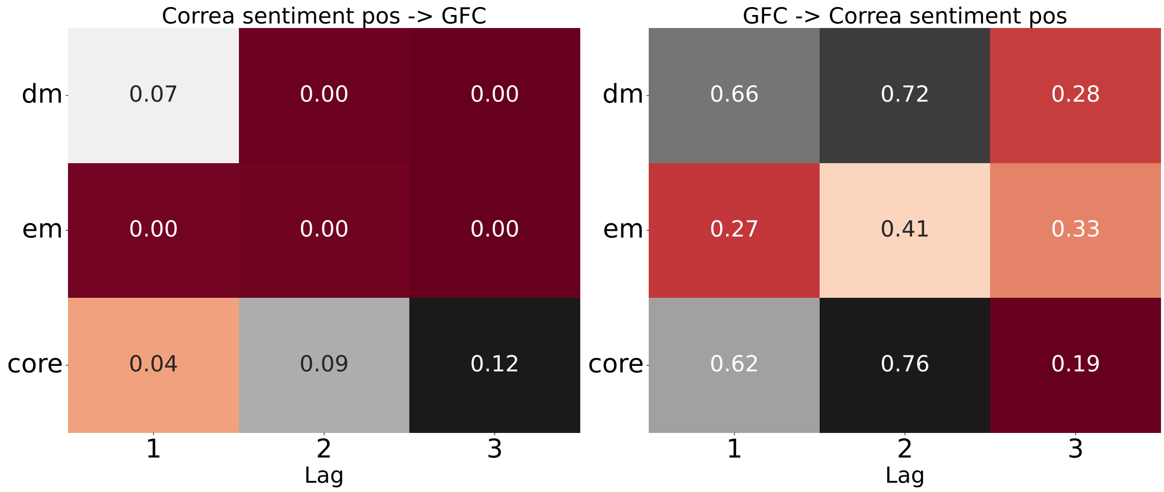
4.7.3 Empirical Evidence for Sentiment Driving the GFC

Beyond descriptive analysis, we find empirical evidence that dictionary-based sentiment expressed in central bank communications may drive components of the GFC. Figure 13 presents results from Granger causality tests using the benchmark Correa sentiment index, showing that aggregated sentiment from both developed and emerging markets Granger-causes movements in the GFC (measured as a 4-quarter difference). Importantly, we do not find evidence of reverse causality, suggesting that central bank communications may play a causal role in shaping global financial conditions. For this analysis, we classify countries following World Bank income categories: high-income economies (per capita GNI above $13,845 in 2022, following the WDI/Atlas-method cutoff) are designated as developed markets, while upper-middle and lower-middle income economies are designated as emerging markets. A complete mapping of each central bank to its classification category is provided in the accompanying metadata file.

Notably, emerging market (EM) sentiment appears to be a more significant driver of the GFC than sentiment from core central banks (ECB, Australia, Canada, Switzerland, UK, Japan, and US). This finding may reflect that EM sentiment captures global investors' risk-on/risk-off preferences, which themselves influence the GFC. Thus, the relationship between EM central bank sentiment and the GFC likely operates through the former's capacity to reflect and amplify global investor risk appetite, rather than through direct causation. These findings extend the literature on international policy transmission by identifying communication as a significant channel through which central banks influence global financial markets.

To assess whether this predictive relationship extends beyond dictionary-based measures, Appendix Figure 19 reports analogous Granger causality tests using LLM-derived, ABS indicators for economic growth and financial conditions. Consistent with the benchmark results based on the Correa index, we find strong evidence that LLM-based sentiment Granger-causes movements in the GFC across multiple model architectures and lag specifications. In particular, sentiment related to economic growth exhibits robust predictive power for the GFC at horizons of one to three quarters, while financial-conditions sentiment also shows

Figure 13: Dictionary-Based Sentiment Granger-Causes the GFC



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

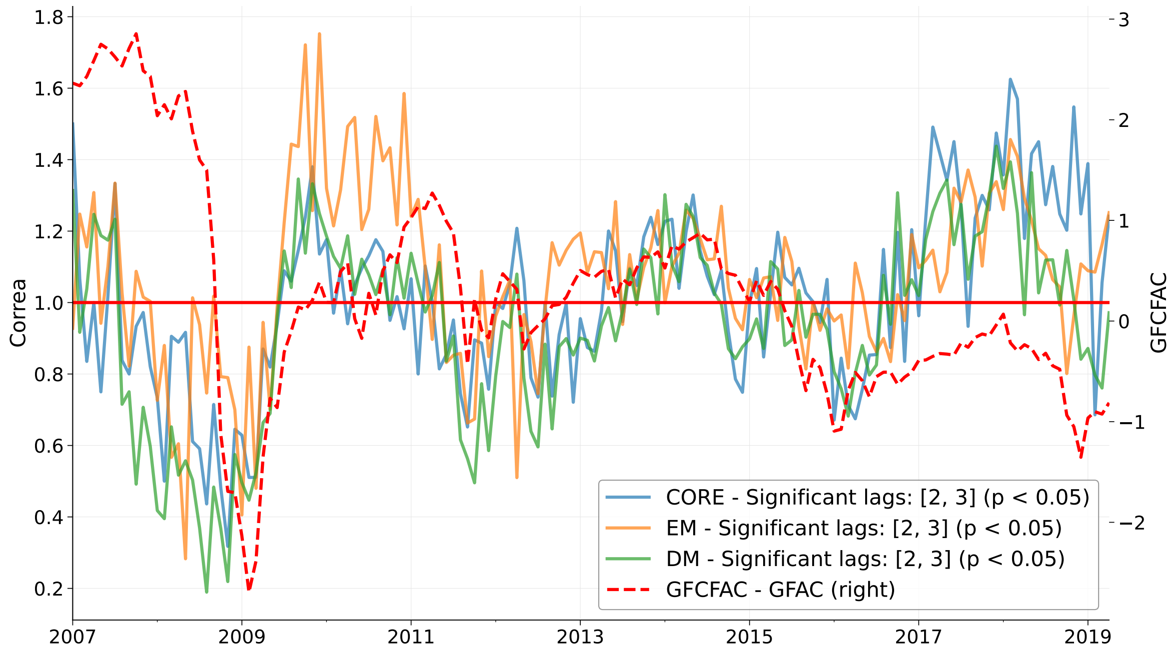
Notes: Sentiment is the benchmark dictionary-based index from Correa et al. (2021) computed from MPSD statements. The GFC is from Miranda-Agrippino and Rey (2020). The figure reports Granger-causality test p-values (rounded to two decimal places) for lag orders $p \in \{1, 2, 3\}$ (see text). "core" denotes a subset of advanced-economy central banks comprising the ECB and the central banks of Australia, Canada, Switzerland, the United Kingdom, Japan, and the United States. "dm" (developed markets) includes Australia, Canada, Switzerland, the Czech Republic, the United Kingdom, Hungary, Israel, Iceland, Japan, South Korea, Norway, New Zealand, Poland, Sweden, Taiwan, and the United States. "em" (emerging markets) includes Argentina, Brazil, Chile, Colombia, Egypt, Ghana, Gambia, Indonesia, India, Kenya, Kazakhstan, Sri Lanka, Liberia, Lesotho, Moldova, Mongolia, Malawi, Mexico, Malaysia, Mozambique, Nigeria, the Philippines, Pakistan, Rwanda, Sierra Leone, Eswatini, Thailand, Tajikistan, Turkey, Ukraine, Uganda, Uzbekistan, South Africa, and Zambia. Country groupings follow World Bank income classifications: high-income economies are classified as developed markets, while upper-middle and lower-middle income economies are classified as emerging markets.

significant predictive content, albeit with greater heterogeneity across models. Importantly, mirroring the main results, we again find little evidence of reverse causality from the GFC to LLM-based sentiment. This asymmetry suggests that LLM-derived sentiment captures forward-looking information embedded in central bank communications, rather than merely responding mechanically to contemporaneous global financial conditions. Taken together, these appendix results reinforce the interpretation that the predictive content of central bank communication for the GFC is not an artifact of a particular sentiment construction, but persists across both traditional lexical indices and modern LLM-based approaches.

4.7.4 Regional Variation in Sentiment

Our analysis also shows important regional variations in how central banks perceive and respond to global financial conditions. Figure 14 illustrates how sentiment evolves differently across major economic regions, potentially reflecting heterogeneous exposure to global shocks and different policy responses.

Figure 14: Regional Sentiment Evolution



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD) and using Correa dictionary.

Notes: Regional sentiment and Global Factor over time. Aggregation: Monthly Mean"Core" denotes a subset of advanced-economy central banks comprising the ECB and the central banks of Australia, Canada, Switzerland, the United Kingdom, Japan, and the United States. "DM" (developed markets) includes Australia, Canada, Switzerland, the Czech Republic, the United Kingdom, Hungary, Israel, Iceland, Japan, South Korea, Norway, New Zealand, Poland, Sweden, Taiwan, and the United States. "EM" (emerging markets) includes Argentina, Brazil, Chile, Colombia, Egypt, Ghana, Gambia, Indonesia, India, Kenya, Kazakhstan, Sri Lanka, Liberia, Lesotho, Moldova, Mongolia, Malawi, Mexico, Malaysia, Mozambique, Nigeria, the Philippines, Pakistan, Rwanda, Sierra Leone, Eswatini, Thailand, Tajikistan, Turkey, Ukraine, Uganda, Uzbekistan, South Africa, and Zambia. Country groupings follow World Bank income classifications: high-income economies are classified as developed markets, while upper-middle and lower-middle income economies are classified as emerging markets.

Figure 14 displays distinct regional patterns, with periods of both convergence and divergence in sentiment. While advanced economies often show similar sentiment trends, emerging markets display more idiosyncratic patterns, particularly during periods of financial stress. These regional differences suggest that the GFC affects economies differently based on their level of financial integration, institutional characteristics, and economic fundamentals.

These findings contribute to the literature on the GFC by providing systematic evidence on how central banks perceive, discuss, and potentially influence global financial conditions through their communications.

5 Discussion

5.1 Policy Implications

Our findings have several implications for monetary policy implementation and coordination. First, the evidence that central bank sentiment Granger causes movements in the GFC (see Figures 13 and 19) suggests that communication strategies may have cross-border spillover effects beyond what is typically acknowledged. This raises questions regarding central banks' responsibility to consider the international implications of their communications, not just their policy actions, and implies that central banks should carefully calibrate their statements, recognizing that they may influence financial conditions well beyond their jurisdictions.

Second, the finding that central banks more frequently cite broad U.S. economic conditions rather than the Federal Reserve specifically as driving global conditions reveals a potential disconnect in how policy spillovers are conceptualized. While the academic literature often focuses on monetary policy transmission across borders (Miranda-Agrippino and Rey, 2020), central bankers themselves appear to emphasize broader economic linkages, suggesting that international policy coordination may need to encompass fiscal and regulatory dimensions as well. The distinction between attributing global influences to a country's economy versus its central bank has meaningful implications for how we model and address international policy spillovers.

Third, the observed regional variations in how sentiment responds to global conditions (as illustrated in Figure 14) highlight the potential for heterogeneous impacts of global shocks, which complicates the case for uniform policy approaches across different economies. This supports arguments for tailored policy frameworks that account for specific vulnerabilities and transmission channels in different economic contexts (Armelius et al., 2020). The substantial differences in how emerging market and advanced economy central banks discuss global conditions suggest that policy coordination efforts should account for these varying perspectives and constraints.

Our linguistic analysis of monetary policy statements (see Table 5 and Figure 6) also reveals how central banks have institutionalized certain communication patterns to convey policy intentions. These formalized linguistic structures may enhance the predictability of policy communications but could potentially constrain adaptability during rapidly evolving economic conditions. Policymakers might benefit from considering how these established communication patterns influence market interpretations and whether greater flexibility in communication styles could improve policy effectiveness in certain contexts.

Beyond these specific findings, our analysis suggests that if central bank narratives actively shape global financial conditions—as our Granger causality results suggest—then coordinating communication strategies may be as important as coordinating policy rate decisions. The post-GFC trend toward longer, more detailed statements reflects a genuine tradeoff: while elaboration enhances transparency, it also increases the complexity of the signal that markets must interpret. Our readability analysis suggests that central banks have partially offset this complexity with simpler sentence structures, but the optimal balance between completeness and clarity remains an open question. The substantial cross-model agreement we document among LLMs suggests that these tools can reliably extract economically meaningful signals from text, but the systematic differences we observe (such as Claude's tendency to identify the Federal Reserve specifically while other models reference the U.S. economy more broadly) highlight the importance of methodological transparency when using these tools for policy analysis.

5.2 The MPSD as a Research Tool

The MPSD offers significant advantages for researchers in economics, finance, and related fields. By providing standardized access to monetary policy communications across multiple countries and time periods, it enables novel comparative analyses previously infeasible due to data limitations. For researchers interested in comparative economics, the database facilitates studies of how institutional differences across central banks manifest in their communication strategies. For those studying policy transmission, the coverage allows for analysis of how market reactions to policy communications vary across different economic environments and institutional frameworks. The significant variations in statement frequency, length, and readability documented in Tables 3 and 4 provide a rich set of institutional variations that can be exploited to identify the determinants and effects of different communication approaches.

The integrated technical infrastructure supporting the database (illustrated in Tables 1 and 2) is particularly valuable for researchers applying computational methods to economic questions. The modular design allows for customization to specific research needs while maintaining reproducibility, addressing a key challenge in computational social science (Gentzkow et al., 2019). Our standardized folder structure (Figure 1) and systematic approach to data processing ensure that analyses can be precisely replicated across different research contexts. The evolution of monetary policy statement characteristics over time, as documented in Figure 5, provides valuable insights into how central bank communication strategies have adapted to changing economic conditions and policy frameworks. Researchers can exploit this temporal variation to study how communication evolves in response to institutional learning, changes in monetary policy regimes, and shifts in the broader economic environment.

5.3 Future Extensions

Several promising directions exist for extending both the database and its applications. First, expanding the corpus to include speeches, minutes, and press conferences would provide a more complete view of central bank communication strategies and enable analysis of how different channels interact to shape market expectations.

Second, incorporating multimedia elements such as press conference transcripts and videos could enable multimodal analysis, including non-verbal cues that may convey policy signals. Advances in computer vision and audio processing make such analysis increasingly feasible.

Third, developing domain-specific pre-training approaches for language models—including continued pre-training on central bank corpora or specialized architectures that account for the hierarchical and temporal structure of policy communications—could improve performance on monetary policy tasks. Our analysis of part-of-speech patterns (Table 5 and Figure 6) provides a foundation for such specialized models.

Finally, integrating the database with high-frequency financial data could facilitate more granular analyses of market reactions to specific communication elements, advancing our understanding of the information effects of monetary policy communications (Nakamura and Steinsson, 2018; Bauer and Swanson, 2023). Building on our finding that sentiment Granger causes the GFC (Figures 13 and 19), such integration would enable more precise identification of which aspects of communication drive market responses across different asset classes.

6 Conclusion

This paper introduces the MPSD, a research infrastructure that addresses fundamental gaps in the empirical analysis of monetary policy communication. By compiling and standardizing 6,693 monetary policy statements from 51 central banks spanning 1990–2024, the database constitutes an unprecedented resource for systematic cross-country analysis, offering exceptional temporal and geographic coverage (Figures 2 and 4). Its technical architecture integrates version control, automated data collection, and standardized preprocessing pipelines (Table 1), ensuring precise replicability while leveraging advanced computational tools. The modular design facilitates extension to diverse research questions (Table 2), supporting the growing integration of NLP methods into economic analysis.

Our linguistic analysis uncovers systematic patterns in how central banks structure their communications (Table 5 and Figure 6), advancing the understanding of institutionalized discourse in monetary policy and informing the development of specialized NLP tools for this domain. Exploiting the breadth of the database, we document a fundamental transformation in central bank communication: statements have approximately tripled in length since 1990, with the expansion accelerating markedly after the Global Financial Crisis, while readability has modestly improved as central banks offset greater informational content with simpler sentence structures (Figure 5). In addition, inflation references comove substantially across countries during episodes of elevated global inflationary pressure, with persistent cross-country differences reflecting heterogeneity in mandates, inflation dynamics, and policy frameworks (Figure 9).

Our empirical application investigates the communication channel of international policy transmission, focusing on the GFC. Using LLMs for question answering and aspect-based sentiment analysis, we document that central banks typically attribute global financial conditions to broad U.S. economic factors rather than specifically to Federal Reserve policy actions. Sentiment analysis reveals greater concern about growth and financial conditions than about monetary policy spillovers in the post-2008 period. Granger causality tests based on both a benchmark dictionary-based sentiment index and LLM-based indicators (Figures 13 and 19) suggest that statement sentiment may drive components of the GFC rather than merely respond to them.

Regional variations in central bank sentiment (Figure 14) reveal important heterogeneity in how central banks perceive and react to global financial conditions, with emerging-market central banks showing greater sensitivity to capital flow concerns and exchange rate pressures relative to their advanced-economy counterparts. These differences indicate that the transmission of global financial shocks depends significantly on institutional factors and the degree of financial integration, carrying important implications for international policy coordination. The comparative analysis of LLM-generated and lexical sentiment indices (Figure 12) illustrates both the potential and the limitations of advanced language models for economic research, underscoring the value of methodological pluralism in computational text analysis.

These findings contribute to the understanding of monetary policy transmission in an interconnected global economy and underscore the importance of the communication channel as a mechanism through which monetary policy transmits across borders. The systematic approach to text analysis developed here offers a replicable template for applying frontier NLP methods to economic research. By making the MPSD and all associated code publicly available, we aim to support future research on monetary policy communication, financial market reactions, and policy coordination. As central banks continue to refine their communication strategies, rigorous empirical analysis of these communications will remain essential for understanding how policy intentions translate into economic outcomes.

7 Acknowledgments

We are grateful to two anonymous referees and the Editor for their constructive comments and suggestions, which substantially improved this paper. We also thank participants at the 4th IFC and Bank of Italy Workshop on Data Science in Central Banking for helpful discussions. All remaining errors are our own.

8 Code and Data Availability

The full codebase used for data processing, analysis, and figure generation, together with the compiled Monetary Policy Statement Database, excluding the scraping code,[5] is publicly available at `https://github.com/CentralBankTexts/monetary-policy-statement-database`. The repository is released under the MIT License, allowing broad reuse while ensuring proper attribution. A permanent archived version of the repository, to ensure long-term accessibility and citability consistent with standard reproducibility practices, will be archived with a permanent DOI upon final publication.

A Appendix

A.1 LLM Sentiment Prompt

```
instruction: |
  You will be analyzing a text from a central bank. The text is provided
   ↪  below:

  <central_bank_text>
  {{CENTRAL_BANK_TEXT}}
  </central_bank_text>

  Your task is to identify discussions related to the impact of global
   ↪  financial conditions on domestic economies. Focus on the following
   ↪  themes:
  1. Economic growth
  2. Financial conditions
  3. Capital flows
  4. Monetary policy spillovers

  Carefully read through the text. For each theme explicitly discussed,
   ↪  assign a continuous sentiment score on a scale from 0.0 to 2.0.
```

[5] While institution-specific scrapers are not distributed, the compiled corpus and full analysis code enable reproduction of all reported results.

```
  Scoring Guidelines:
  * 0.0 to < 1.0 (Negative): The text suggests unfavorable impacts, risks,
   ↪  contraction, or tightening.
      * Closer to 0.0: Severe risks, crises, or strong negative shocks.
      * Closer to 1.0: Mild headwinds or slight downside risks.
  * 1.0 (Neutral): The text presents a balanced view, negligible impact,
   ↪  or uncertainty with no clear direction.
  * > 1.0 to 2.0 (Positive): The text suggests favorable impacts,
   ↪  expansion, loosening, or resilience.
      * Closer to 1.0: Mild tailwinds or slight improvements.
      * Closer to 2.0: Strong growth, robustness, or highly beneficial
       ↪  conditions.

  Output Rules:
  * Only include themes that are explicitly discussed.
  * Use floating-point numbers to capture nuance (e.g., 0.45, 1.2, 1.85).
  * Format the output as a Python list of lists, where each inner list
   ↪  contains the theme string and the score.

  Your final output should be a Python list containing only the
   ↪  theme-score pairs. Do not provide any additional commentary.

  Example Output:
  [["economic_growth", 1.8], ["financial_conditions", 0.45],
   ↪  ["capital_flows", 1.0]]

output_format: |
  A Python list of lists containing theme and float score pairs (e.g.,
   ↪  [["economic_growth", 1.25], ["financial_conditions", 0.8]]).
```

A.2 LLM Narrative Prompt

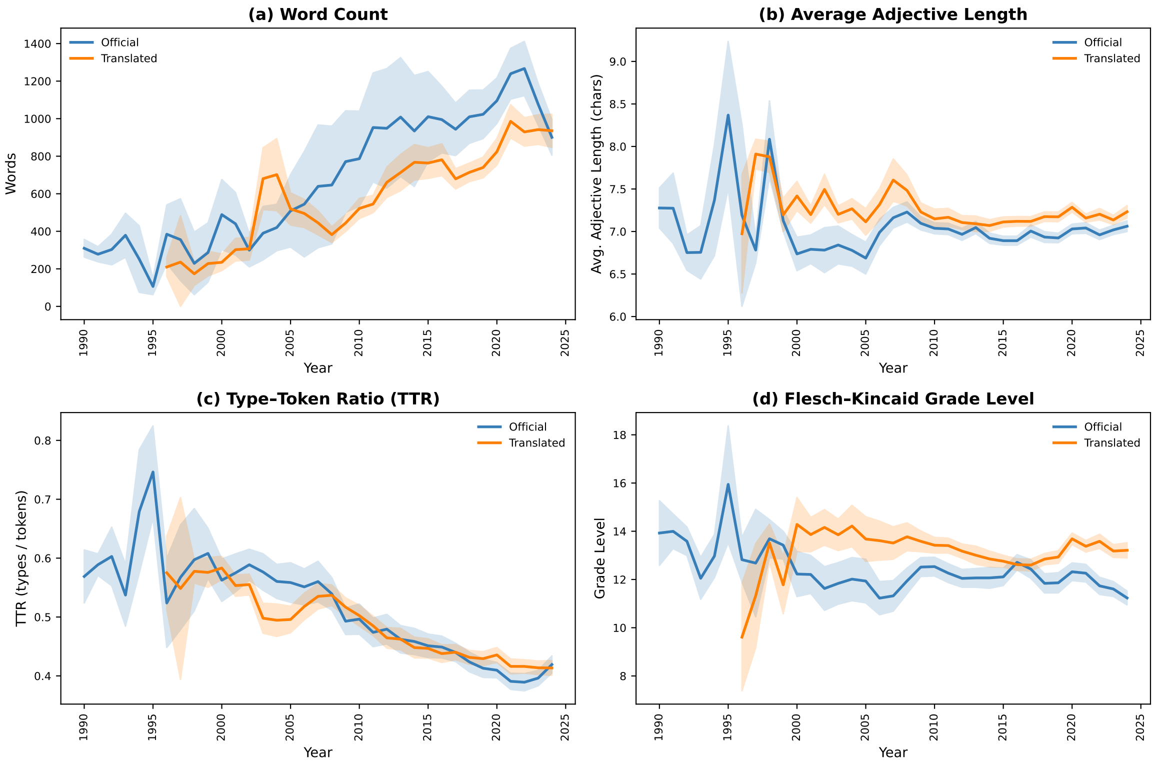
```
instruction: |
  The following texts come from central banks. The country is identified
   ↪  in the Country: and the Comment: is the text released by the central
   ↪  bank.
  I want you to identify the main driver of global economic and financial
   ↪  conditions and the origin country driving these conditions.
  In the summary please note if any central banks identify the Federal
   ↪  Reserve as the driver of global economic or financial conditions.

output_format: |
  Please respond in the following format:
  Driver: [Identified Driver]
  Origin Country: [Origin Country]
```

```
Summary: [brief explanation]
```

A.3 Language Groups Analysis

We distinguish between monetary policy statements originally drafted in English and those translated into English from other languages. Figure 15 reports the evolution of linguistic indicators separately for these two groups, allowing us to assess whether observed patterns are driven by language or translation effects. For statements originally drafted in English, the results closely mirror the full-sample trends. Average word counts increase markedly over time, while lexical diversity, as measured by the type-token ratio, declines. At the same time, readability exhibits only a modest downward trend, indicating that although statements have become substantially longer, they have not become proportionally more complex in terms of sentence structure and word choice. Statements translated into English display broadly similar dynamics. Despite differences in levels, the time-series trends closely track those observed for native-English statements: translated texts become longer over time, show declining lexical diversity, and maintain relatively stable readability. Average adjective length remains largely unchanged in both groups, suggesting that the expansion of communication occurs primarily through increased volume rather than more complex evaluative wording. The similarity of trends across native-English and translated-English statements suggests that the main results are not driven by translation-related measurement artifacts. Instead, they reflect systematic changes in central bank communication practices over time. Persistent level differences across groups are more plausibly attributed to institutional drafting conventions and communication styles rather than language effects per se.

Figure 15: Average Statement Length and Readability Measure by Language Groups



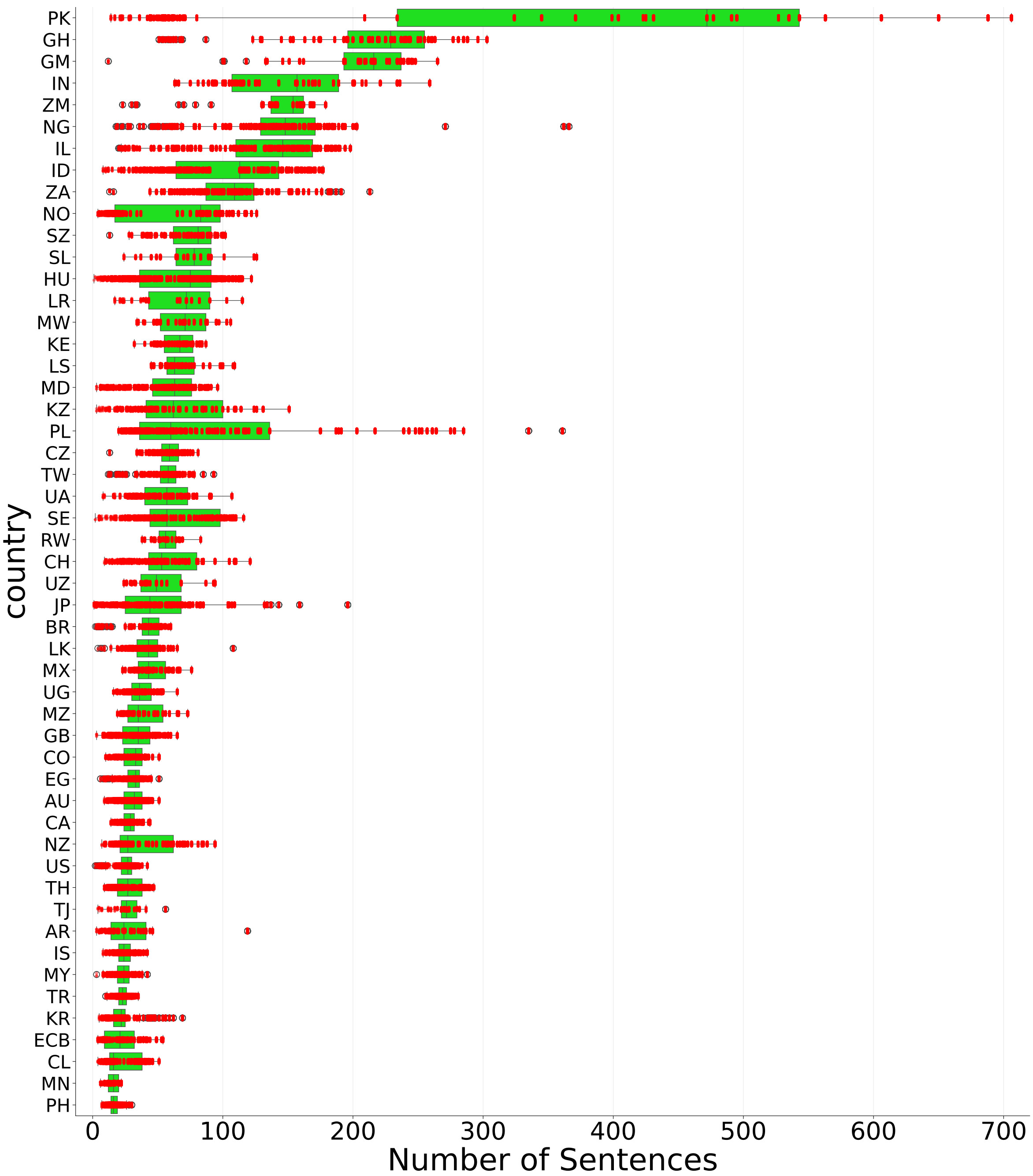
Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

Notes: For each central bank statement, we compute four document-level linguistic indicators for countries with English as official language and English as translated language: (i) word count, defined as the total number of tokens; (ii) average adjective length, measured as the mean character length of adjectival tokens identified using part-of-speech tagging; (iii) the type–token ratio (TTR), defined as the number of unique tokens divided by total tokens; and (iv) readability, captured using standard readability indices (e.g., Flesch-Kincaid Grade Level). Each indicator is first computed at the statement level and then aggregated to the annual frequency by taking the cross-statement mean. To capture uncertainty and heterogeneity across institutions and over time, we report confidence bands constructed from the standard error of the annual mean.

A.4 Statement Sentence Count

We observe that unusually long "statements" in the corpus are concentrated in countries where the policy-rate decision is communicated through a broader analytical document rather than a short press release (see Figure 16). For example, in Pakistan (PK) the rate decision has often been published within, or together with, a Monetary Policy Statement/Report that includes detailed discussion of global and domestic economic developments and the outlook. A similar pattern appears in several other countries as well, such as Nigeria (NG) or Ghana (GH), for example, where MPC statements typically combine the decision with an extended review of macroeconomic conditions and risks. In India (IN), statement length is amplified by the release of a policy package around each meeting (e.g., MPC resolution together with accompanying policy statements), which can be captured as a single text. Large sentence-count variation primarily reflects institutional differences in communication formats and document bundling.

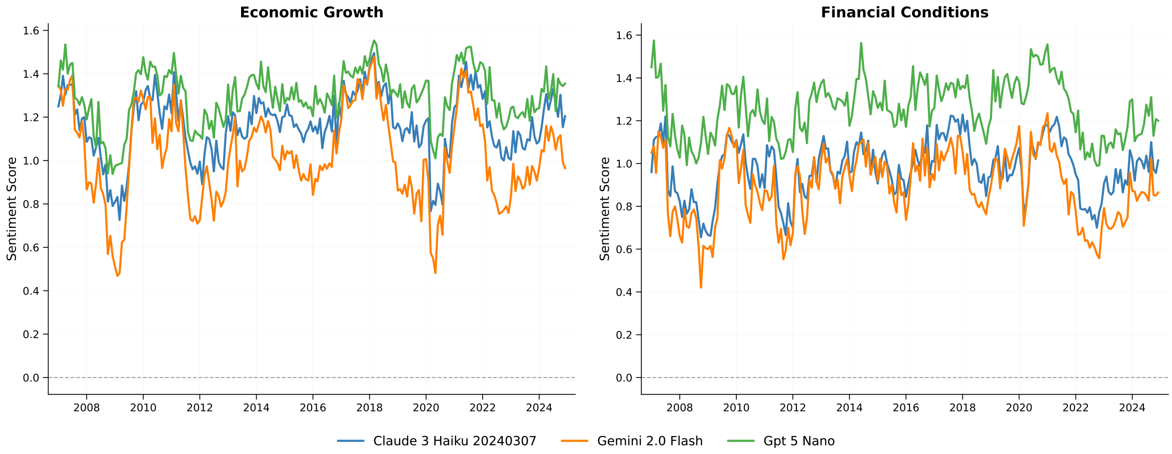
Figure 16: Statement Sentence Count



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD) and using LLMs.

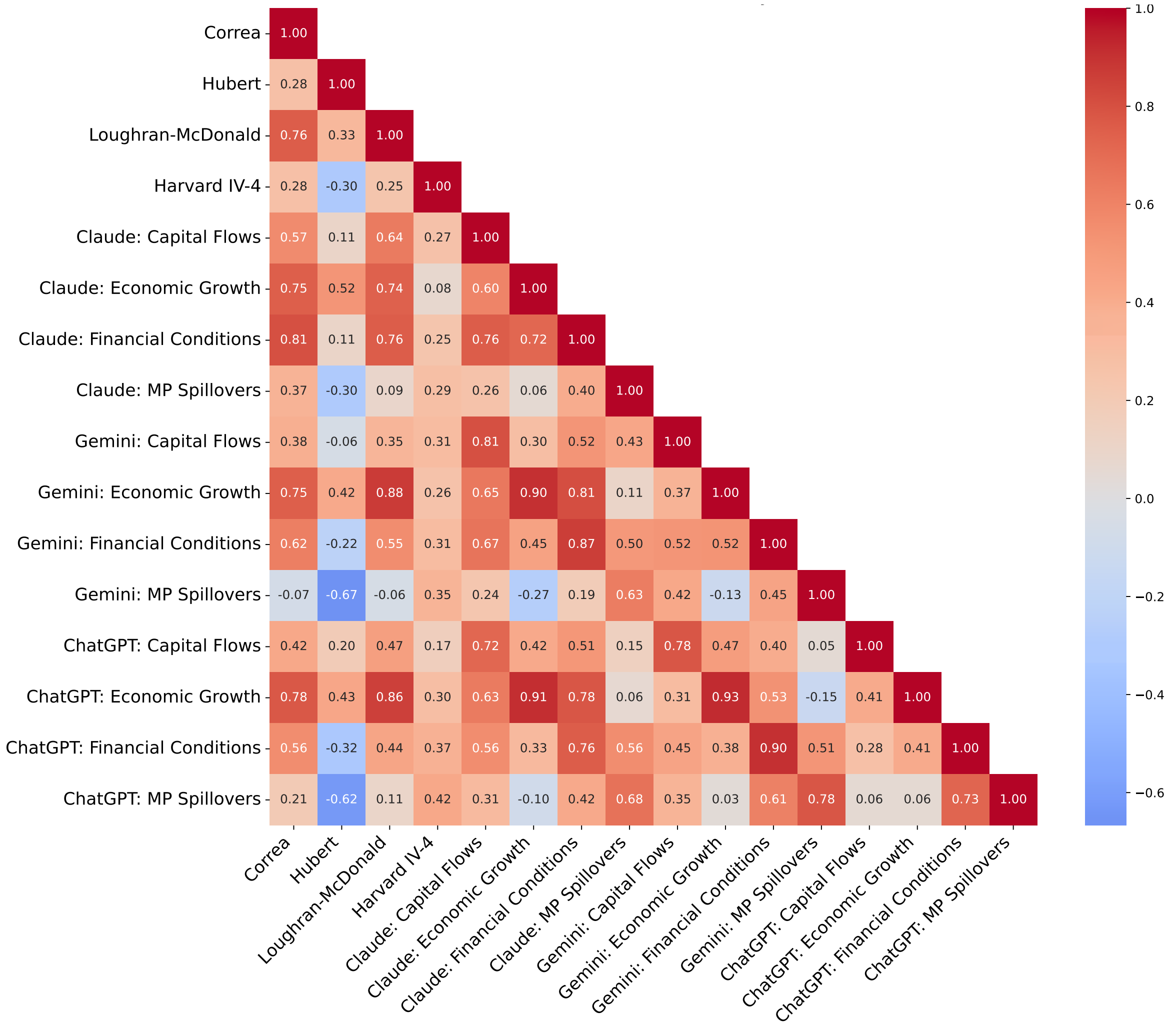
Notes: Statement length in sentences by country.

Figure 17: Cross-Model LLM Sentiment Comparison



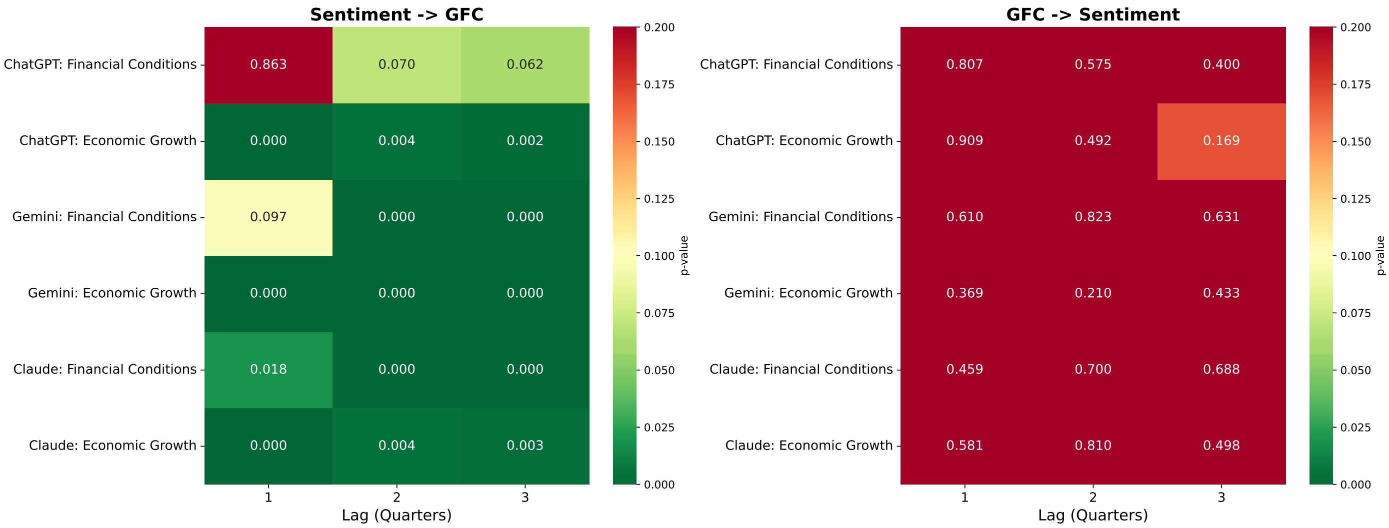
Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD) and using LLMs.

Figure 18: Extended Correlation Matrix of Sentiment Measures



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD) and using LLMs.

Figure 19: LLM-Based Sentiment Granger-Causes the GFC



Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

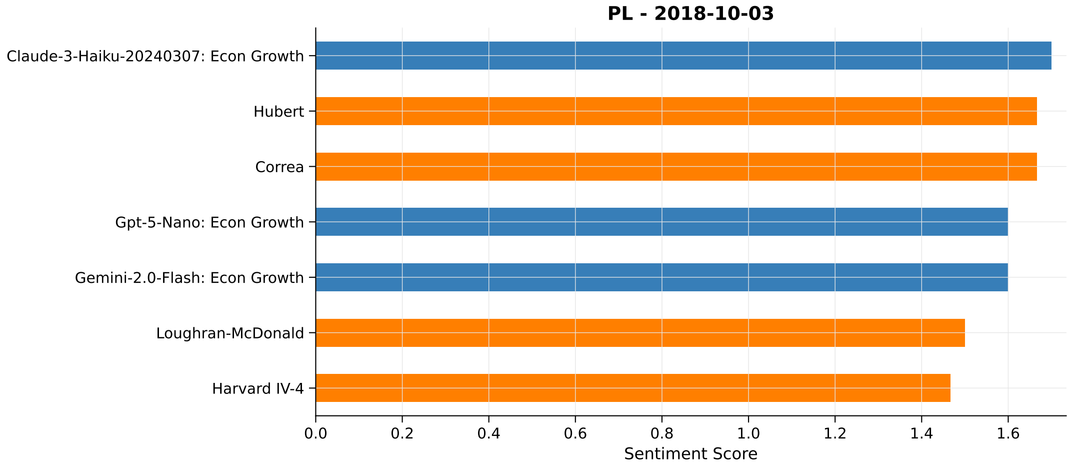
Notes: Sentiment measures are derived from MPSD statements using multiple LLMs (ChatGPT, Gemini, and Claude) and capture ABS related to economic growth and financial conditions. The GFC is from Miranda-Agrippino and Rey (2020). The figure reports p-values (rounded to three decimal places) from Granger causality tests at quarterly frequency. Green cells indicate statistical significance at the 5% level. Tests are conducted for lags of one to three quarters.

Figure 20: Most Common Three-Token Part-of-Speech Pattern `DET PROPN VERB`.

Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

Notes: The `DET PROPN VERB` pattern often forms the root of sentences associated with CB. Total count of three-token sentences: 6,685.

Figure 21: Sentiment Variation Across Example Statement, Poland 2018-10-03



Warsaw, 3 October 2018 Information from the meeting of the Monetary Policy Council held on 2-3 October 2018 The Council decided to keep the NBP interest rates unchanged: reference rate at 1.50%; lombard rate at 2.50%; rediscount rate at 0.50%; rediscount rate at 1.75%. Global economic conditions remain favourable. In the euro area, the economic situation continues to be favourable, despite slightly slower GDP growth than in 2017. In the United States, economic growth is higher than in the previous year, which confirms that economic conditions in this country are strong. In China, activity growth has continued at a stable pace for the past few quarters. Since the beginning of the year, global energy commodity prices, including those of oil, have risen substantially. This has contributed to higher inflation in many countries. At the same time, core inflation in the external environment of the Polish economy, including the euro area, remains moderate, despite continued strong econ

Source: Authors' calculations based on the Monetary Policy Statement Database (MPSD).

Notes: This figure illustrates the qualitative validation approach by displaying sentiment variation across methods for the same policy statement

References

Acosta, M. and E. E. Meade (2015). Hanging on every word: semantic analysis of the FOMC's postmeeting statement. FEDS Notes 2015-09-30, Board of Governors of the Federal Reserve System.

Armelius, H., C. Bertsch, I. Hull, and X. Zhang (2020). Spread the Word: International spillovers from central bank communication. *Journal of International Money and Finance 103*, 102116.

Assenmacher, K., G. Glöckler, S. Holton, P. Trautmann, D. Ioannou, S. Mee, K. Bakk-Simon, S. Bergbauer, M. Catenaro, E. Charalampakis, M. Ehrmann, G. Ferrero, D. Georgarakos, P. Gertler, A. Giovannini, R. Grandia, N. Hernborg, N. Herrala, D. Kedan, G. Kenny, T. Linzert, M. Manrique, R. Mestre, E. Mönch, S. Nardelli, N. Nomm, L. Pavlova, G. Schultefrankenfeld, M. Silgoner, I. Skotida, B. Winkler, M. T. Bitterlich, E. Argiri, C. Alonso, J. Byron, F. Arigoni, M. Deroose, M. Gardt, K. Istrefi, O. Goldfayn-Frank, J. Hellström, P. Huertgen, K. Kalnberzina, G. Kocharkov, V. Márquez, C. Ruhe, M. Šanta, R. Reedik, G. Sciot, A. Stylianou, and E. Taylor (2021). Clear, consistent and engaging: ECB monetary policy communication in a changing world. Technical Report 274, ECB Occasional Paper.

Bauer, M. D. and E. T. Swanson (2023). An alternative explanation for the "Fed Information Effect". *American Economic Review 113*(3), 664–700.

Benchimol, J., I. Caspi, and S. Kazinnik (2023). Measuring communication quality of interest rate announcements. *Economists' Voice 20*(1), 43–53.

Benchimol, J., S. Kazinnik, and Y. Saadon (2020). Communication and transparency through central bank texts. Paper presented at the 132nd Annual Meeting of the American Economic Association, January 3-5, 2020, San Diego, CA, United States.

Benchimol, J., S. Kazinnik, and Y. Saadon (2022). Text mining methodologies with R: an application to central bank texts. *Machine Learning with Applications 8*, 100286.

Benchimol, J., S. Kazinnik, and Y. Saadon (2025). Federal Reserve Communication and the COVID-19 Pandemic. *Manchester School 93*(5), 464–484.

Blinder, A. (2009). Talking about monetary policy: the virtues (and vice?) of central bank communication. BIS Working Papers 274, Bank for International Settlements.

Blinder, A., C. Goodhart, P. Hildebrand, D. Lipton, and C. Wyplosz (2001). *How Do Central Banks Talk?* Number 3 in Geneva Reports on the World Economy. London, UK: Centre for Economic Policy Research.

Blinder, A. S. (2018). Through a crystal ball darkly: The future of monetary policy communication. *AEA Papers and Proceedings 108*, 567–571.

Blinder, A. S., M. Ehrmann, J. de Haan, and D.-J. Jansen (2024). Central bank communication with the general public: promise or false hope? *Journal of Economic Literature 62*(2), 425–457.

Blinder, A. S., M. Ehrmann, M. Fratzscher, J. D. Haan, and D.-J. Jansen (2008). Central bank communication and monetary policy: a survey of theory and evidence. *Journal of Economic Literature 46*(4), 910–945.

Born, B., M. Ehrmann, and M. Fratzscher (2014). Central bank communication on financial stability. *Economic Journal 124*(577), 701–734.

Bruno, V. and H. S. Shin (2015). Capital flows and the risk-taking channel of monetary policy. *Journal of Monetary Economics 71*, 119–132.

Brusa, F., P. Savor, and M. Wilson (2020). One Central Bank to Rule Them All. *Review of Finance 24*(2), 263–304.

Bulír, A., M. Cihák, and D.-J. Jansen (2013). What drives clarity of central bank communication about inflation? *Open Economies Review 24*(1), 125–145.

Campiglio, E., J. Deyris, D. Romelli, and G. Scalisi (2025). Warning words in a warming world: Central bank communication and climate change. *European Economic Review 178*, 105101.

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

Coibion, O., Y. Gorodnichenko, and M. Weber (2022). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy 130*(6), 1537–1584.

Correa, R., K. Garud, J. M. Londono, and N. Mislang (2021). Sentiment in central banks' financial stability reports. *Review of Finance 25*(1), 85–120.

Davis, J. S. and M. A. Wynne (2019). Central Bank Communications: A Case Study. In *Oxford Handbook of the Economics of Central Banking*. Oxford University Press.

De Grauwe, P. and Y. Ji (2023). On the use of current and forward-looking data in monetary policy: a behavioural macroeconomic approach. *Oxford Economic Papers 75*(2), 526–552.

de Haan, J. and L. Hoogduin (2024). ECB communication policies: An overview and comparison with the Federal Reserve. *Journal of International Money and Finance 142*, 103050.

Dincer, N., B. Eichengreen, and P. Geraats (2022). Trends in monetary policy transparency: further updates. *International Journal of Central Banking 18*(1), 331–348.

Ehrmann, M. and A. Wabitsch (2022). Central bank communication with non-experts - A road to nowhere? *Journal of Monetary Economics 127*, 69–85.

Eijffinger, S. and D. Masciandaro (2014). *Modern Monetary Policy and Central Bank Governance*. Cheltenham, UK: Edward Elgar Publishing.

Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as data. *Journal of Economic Literature 57*(3), 535–574.

Gonzalez, M. and R. C. Tadle (2022). Monetary policy press releases: an international comparison. BIS Working Papers 1023, Bank for International Settlements.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37*(3), 424–438.

Hansen, S. and M. McMahon (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics 99*(S1), 114–133.

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

Hayo, B. and M. Neuenkirch (2015). Central bank communication in the financial crisis: Evidence from a survey of financial market participants. *Journal of International Money and Finance 59*(C), 166–181.

Issing, O. (2019). *The long journey of central bank communication*. Cambridge, MA: MIT Press.

Kahn, G. A. (2007). Communicating a policy path: the next frontier in central bank transparency? *Economic Review-Federal Reserve Bank of Kansas City 92*(1), 25.

Kalemli-Özcan, Ş. (2019). U.s. monetary policy and international risk spillovers. In *Proceedings - Jackson Hole Economic Symposium*, Jackson Hole, WY, pp. 95–191.

Kryvtsov, O. and L. Petersen (2021). Central bank communication that works: Lessons from lab experiments. *Journal of Monetary Economics 117*(C), 760–780.

Kumar, A., S. K. Mallick, M. Mohanty, and F. Zampolli (2023). Market volatility, monetary policy and the term premium. *Oxford Bulletin of Economics and Statistics 85*(1), 208–237.

Lewis, D., C. Makridis, and K. Mertens (2020). Do monetary policy announcements shift household expectations? CEPR Discussion Papers 14360, Centre for Economic Policy Research.

Loughran, T. and B. McDonald (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance 66*(1), 35–65.

Macklem, T. and J. Vardy (2023). 20 years of central bank communications, and lessons for the future. In *Central banking in the Americas: Lessons from two decades*, pp. 55–67. Bank for International Settlements.

Masciandaro, D., O. Peia, and D. Romelli (2024). Central bank communication and social media: from silence to Twitter. *Journal of Economic Surveys 38*(2), 365–388.

Miranda-Agrippino, S. and H. Rey (2020). U.S. monetary policy and the Global Financial Cycle. *Review of Economic Studies 87*(6), 2754–2776.

Miranda-Agrippino, S. and H. Rey (2022). The global financial cycle. In *Handbook of International Economics*, Volume 6, pp. 1–43. Elsevier.

Morris, S. and H. S. Shin (2002). Social value of public information. *American Economic Review 92*(5), 1521–1534.

Nakamura, E. and J. Steinsson (2018). High-frequency identification of monetary non-neutrality: the information effect. *Quarterly Journal of Economics 133*(3), 1283–1330.

Obstfeld, M. (2015). Trilemmas and tradeoffs: living with financial globalization. In C. Raddatz, D. Saravia, and J. Ventura (Eds.), *Global Liquidity, Spillovers to Emerging Markets and Policy Responses*, Volume 20, Chapter 2, pp. 013–078. Central Bank of Chile.

Picault, M. and T. Renault (2017). Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance 79*, 136–156.

Rey, H. (2015). Dilemma not trilemma: the global financial cycle and monetary policy independence. NBER Working Papers 21162, National Bureau of Economic Research, Inc.

Vyshnevskyi, I., W. Jombo, and W. Sohn (2024). The clarity of monetary policy communication and financial market volatility in developing economies. *Emerging Markets Review 59*, 101121.

Woodford, M. (2005). Central bank communication and policy effectiveness. *Proceedings - Economic Policy Symposium - Jackson Hole Aug*, 399–474.

The Monetary Policy Statement Database

An LLM Application to Global Financial Conditions*

Cory Baird[1] Jonathan Benchimol[2] Wook Sohn[3] Vira Vyshnevska[3] Iegor Vyshnevskyi[4]

[1]University of Tokyo, Tokyo, Japan
Corresponding author. Email: cbaird@g.ecc.u-tokyo.ac.jp

[2]Bank of Israel, Jerusalem, Israel

[3]KDI School, Sejong, South Korea

[4]Sogang University, Seoul, South Korea

February 18, 2025

*Code and data: `github.com/CentralBankTexts/monetary-policy-statement-database`

Outline

Not Just a Database: Reproducible AI for Macroeconomics

- **Scientific Rigor**
  Infrastructure to maintain scientific standards in text analysis.
- **Unprecedented Scale**
  6,693 policy rate statements from 51 central banks worldwide.
- **Cutting-Edge Infrastructure**
  Transparent, automated, and scalable pipelines.
- **Applying AI to Macroeconomics**
  Leverages AI to identify economic narratives and sentiment over time and across countries.
  - Global Financial Cycle origins and drivers (Rey, 2015).

Current Literature

Key Limitations in CB Communication Research:

- **Data & Standardization:** Existing research focuses on a few countries or uses only BIS speeches (Born et al., 2014; Picault and Renault, 2017).
  - Contribution: standardized dataset of 51 CBs
    Larger than existing non-public databases, e.g., Gonzalez and Tadle (2022)
- **Reproducibility:** Lack of open-source pipelines impedes reproducible research (Gentzkow et al., 2019; Hansen and McMahon, 2016).
  - Contribution: transparent and accessible text cleaning, and Large Language Model (LLM) pipelines

Applying the Scientific Method to "AI" and NLP

- Failure to share data/algorithms hinders reproducibility and efficiency
  - **Can reproduce** lexical/dictionary methods at immense cost
  - **Cannot reproduce** black box LLMs (Claude, Gemini) because data, prompts, and models must be **exactly** the same

- **Irreproducible NLP methods** hinder scientific progress and data reliability in downstream modeling

Data Collection	**Featurization**	**Replicable AI Use**
Identify statement links	Cleaning	Replicable LLM infrastructure
Scrape statements	Sentence segmentation	Track inputs (prompts, texts)
Automate retrieval	Part-of-speech tagging	Track outputs (responses)

Table 1: End-to-End Transparency for Replicable AI

Source: Authors.

Infrastructure Overview

- **Transparency**
  - All code/scripts made public on GitHub[1]
  - Technologies: GitHub, Pyenv, Poetry
- **Automation**
  - Scripts executed regularly via a Data Version Control (DVC) pipeline
  - Technologies: DVC, GitHub Actions.
- **Scalability**
  - Modular scripts easily adaptable for other CB texts
  - Technologies: Python Object Oriented Programming, GitHub
- **Reproducibility**
  - Transparent text analysis
  - Reproducible pipelines for cleaning, featurization, and LLMs
  - Technologies: UV, Pyenv, Poetry, S3, DVC

[1] `github.com/CentralBankTexts/monetary-policy-statement-database`

MPSD Overview

- Monetary Policy Statement Database (MPSD)
  - **Scope**
    51 CBs (35 API and 16 static currently cleaned)
  - **Data**
    Approximately 6,693 unique MPSs
  - **Temporal Coverage**
    1990–2024

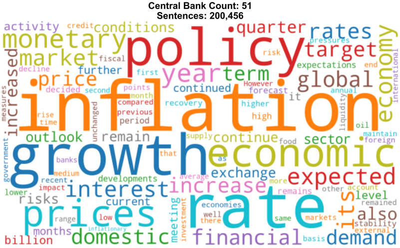


Figure 1: Standardized and diverse corpus for training/testing A.I.

Outline

Technical Infrastructure: Code & Data Management

- **Data integrity and reproducibility**
  - Data Version Control (DVC) tracks full lifecycle: web scraping $\rightarrow$ LLM analysis
  - Parameterized pipelines enforce data lineage at every stage
- **Code and environment management**
  - Code versioning: Git and GitHub
  - Python version and dependency management: uv
  - Cloud storage: AWS S3 (raw, processed, and artifact partitions)
- **Automated updates**
  - GitHub Actions: daily checks for new CB policy statements
  - Single-command replication of data, models, and outputs
- **LLM infrastructure**
  - Abstract LLM class integrating Claude, ChatGPT, Gemini, and Hugging Face
  - Methods: model selection, YAML prompt processing, response generation/saving

Code/Folder Infrastructure: Automated & Scalable

- **Modular Design:** Programs/scripts are modularized for:
  1. Text retrieval
  2. Text cleaning
  3. LLM inference
- **Standardized Folder Structure:**

  `src` : Python scripts
  `data` : Text files
  `references` : GenAI prompts
- **MPSD Updates:** Automated updates using DVC.
- **End-to-End Reproducibility:** A full pipeline ensures reproducible results from data acquisition to NLP analysis.

Outline

The Rise of Central Bank Communication

- CB communication dramatically changed since the 1990s
  - First official MPS: Reserve Bank of Australia in 1990
  - Rapid growth in MPSs after the year 2000
- MPSs are a crucial tool for:
  - Managing inflation expectations
  - Enhancing policy predictability
  - Building credibility and trust

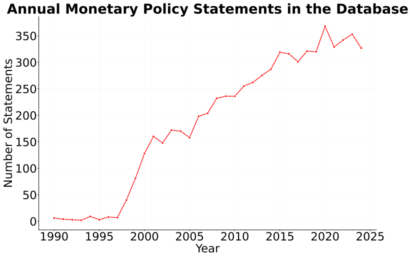


Figure 2: Annual MPSs in the MPSD

Dataset Coverage

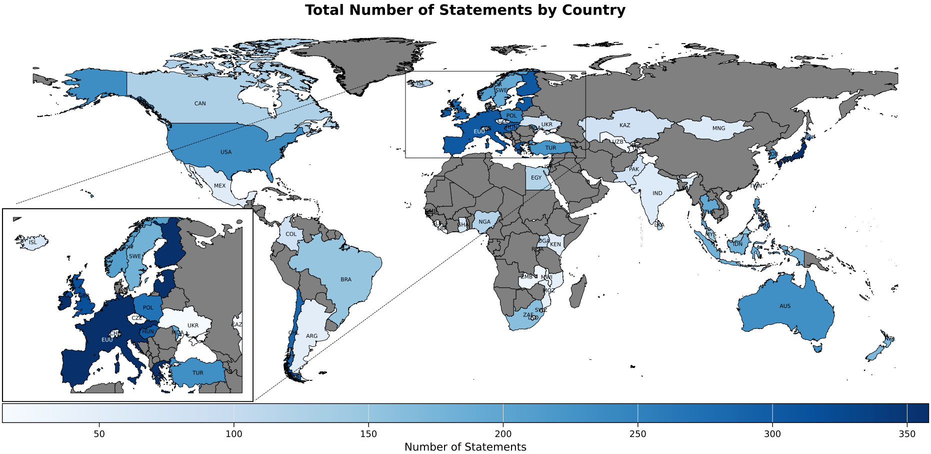


Figure 3: Countries and Monetary Policy Statements

Dataset Coverage

- The BoJ (Japan) and ECB emerge as the most prolific publishers of statements
- Central banks in middle-income countries exhibit more limited publication frequency and shorter historical coverage
- These cross-institutional variations reflect differences in institutional frameworks and communication strategies

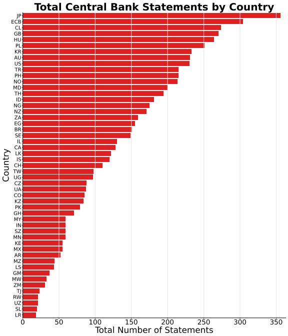


Figure 4: Statements by Country

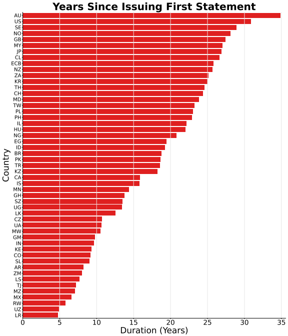


Figure 5: Years of Coverage by Country

Trends in Communication Over Time

- MPSs growing lengthier and more complex post-GFC
- MPSs becoming slightly more "readable" over time despite increased length

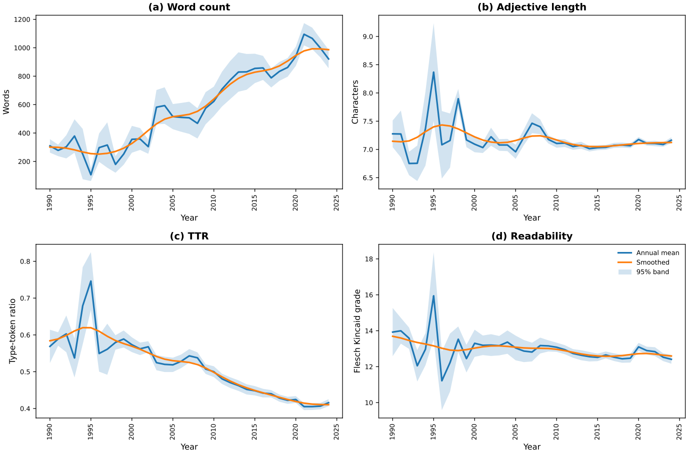


Figure 6: Linguistic Trends in CB Statements Over Time

Key Characteristics (Patterns) of the MPS

Economic Theme	**POS Pattern**	**Example**
Inflation Forecasts	inflation + AUX + VERB	inflation **will rise**; inflation **might persist**
Growth Projections	growth/GDP + AUX + VERB	growth **should recover**; GDP **may stall**
Inflation Dynamics	NOUN + ADP + inflation	risk **of** inflation; outlook **for** inflation
Rate Levels	rate + ADP + NUM	rate **at** 5.25%; rate **above** 2.0
Forward Guidance	policy/rate + AUX + VERB	policy **will remain**; rates **might stay**
Monetary Stance	policy + AUX + ADJ	policy **is restrictive**; stance **remains loose**
Fiscal Policy	NOUN + ADP + spending/deficit	impact **of** deficit; boost **from** spending
External Conditions	NOUN + ADP + trade/export	demand **for** exports; uncertainty **about** trade

Table 2: Linguistic Patterns in Monetary Policy Statements by Economic Theme

Most Common Sentence Patterns by Part-of-speech

- Common sentence patterns include:
  - `DET ADJ NOUN` (e.g., "The strong dollar," "The current rate")
  - `DET NOUN ADP` (e.g., "The impact of," "The level of")
  - `ADP DET NOUN` (e.g., "In the market," "On the economy")
- Sentences identified with spaCy's Universal POS tags and pattern matching.

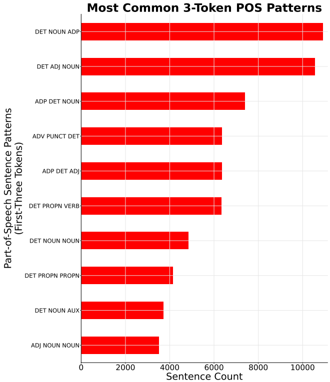


Figure 7: Most Common Patterns by Part-of-Speech (POS)

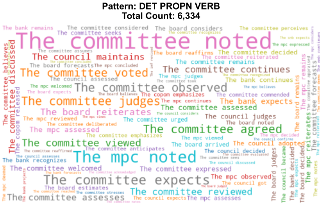


Figure 8: POS: The `DET PROPN VERB` pattern is associated with CB commentary/actions

Outline

Addressing AI Challenges with the Corpus/Infrastructure

- **Challenges with LLMs**
  1. Irreproducibility: Insufficient transparency regarding model inputs and specifications makes replicating it difficult.
     - We provide open-source code, including LLM code, prompts, and Jupyter notebooks, ensuring full transparency and reproducibility.
  2. High computational costs and energy consumption.
     - We implement text filtering (keyword & part-of-speech) pipelines to reduce the computational burden.

- **Application to Econ: The Global Financial Cycle (GFC)**
  - Communication Channel: CB Communication $\Rightarrow$ Market Expectations $\Rightarrow$ Financial Conditions
  - Using our LLM/Data Pipeline we leverage LLMs for Question Answering (QA) & classification (topic, sentiment) tasks to explore origins and drivers of the Global Financial Cycle (Miranda-Agrippino and Rey, 2020, 2022)

Leveraging Our LLM/Data Pipeline

- Infrastructure is abstract for scalability
  - LLM: seamlessly add more models (ChatGPT, Huggingface, etc) and tasks (classification, summarization, Q&A)
  - Text retrieval: seamlessly add more central banks and types of communications (minutes, speeches, etc)

Modular LLM-based Data Processing Framework for Academic Research		
Data Acquisition Module (Abstract Class)	**LLM Processing Module (Abstract Class)**	**LLM Data Pipeline Workflow**
Methods: 1. Initialize (source parameters) 2. Scrape Statement Links 3. Download/Save Statements **Output:** Text Files CSV/Parquet Text Dataframes **Abstraction:** Extend abstract functions to accommodate new data sources (e.g., central bank speeches, minutes) and communication modes.	**Methods:** 1. Initialize (LLM Model) 2. Process/Load Prompts 3. Generate Response **Output:** Response (Text/JSON) Prompt (Text/JSON) Model Parameters (YAML) **Abstraction:** Implement model-agnostic functions to leverage diverse LLM architectures (e.g., Gemini, Claude, HuggingFace models).	**Methods:** 1. Initialize (Input Corpus) 2. Filter via Keywords or POS 3. LLM Inference 4. Save Response **Output:** Response (Text/JSON) Prompt (Text/JSON) Model Parameters (YAML) Figure for Research (PNG/PDF) **Use-Case:** Utilize concise, reusable scripts for seamless integration with LLM APIs (e.g., Gemini, Claude) by researchers.

Table 3: AI Infrastructure for Central Bank Communication Analysis

LLM Application: Exploring the Global Financial Cycle

- LLM Application: Question Answering (QA)
  - Prompts model to summarize drivers of global economy
    "The following texts come from central banks. [. . .] I want you to identify the main driver of global economic and financial conditions and the origin country driving these conditions. In the summary please note if any central banks identify the Federal Reserve as the driver of global economic or financial conditions."
- LLM Application: Aspect Based Sentiment[2](ABS)
  - Prompts model to extract sentiment for specific themes on a continuous scale (0.0–2.0)
    "Your task is to identify discussions related to the impact of global financial conditions on domestic economies. Focus on the following themes: 1. Economic growth, 2. Financial conditions, 3. Capital flows, 4. Monetary policy spillovers. For each theme explicitly discussed, assign a continuous sentiment score on a scale from 0.0 to 2.0, where 1.0 is neutral."
- Findings from LLM Analysis:
  - QA: Central Banks often cite the U.S. economy as a driver of global economic conditions rather than specifically the U.S. Federal Reserve.
  - ABS: Sentiment analysis shows that central banks expressed more concern about growth and financial conditions than monetary policy and capital flows in the post-2008 period

Appendix: Full Prompt Example

[2] Aspect-based sentiment identifies sentiment of specific theme; see Appendix A.1–A.2 in the paper for full prompts

LLM Application: QA Results Differ by Model (2008)

Model	Mechanism	Origin	Key Finding
Claude	The Federal Reserve	United States	Names the Fed as key driver; US housing bust a recurring theme
Xiaomi	Global financial market turmoil / subprime crisis	United States	Fed acknowledged by multiple CBs as key driver
Mistral	US economy slowdown / subprime crisis	United States	No CB explicitly names the Fed as the driver
Arcee	US subprime crisis and broader US economic weakness	United States	Fed's policy actions cited as relevant but not the primary driver

LLM summaries of the same CB sentences differ in how prominently they identify the Federal Reserve

LLM Application: Aspect Based Sentiment

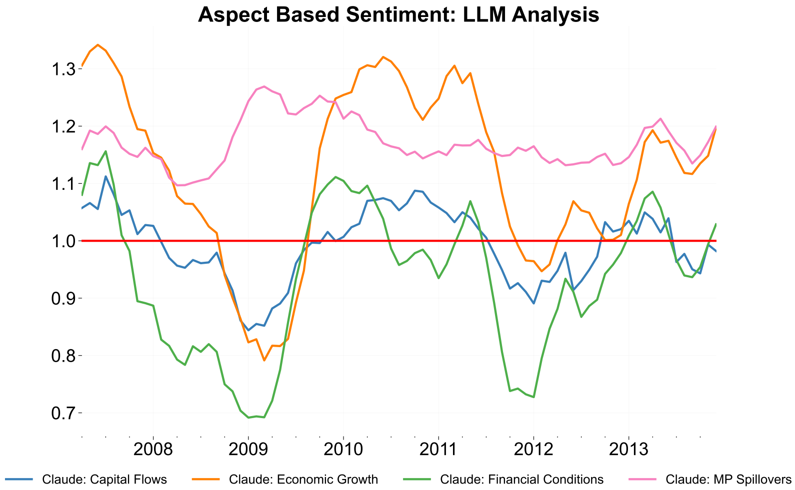


Figure 10: LLM can easily produce aspect based sentiment

Sentiment Validation: LLM Complements and Refines Lexical Methods

- LLM-based growth sentiment: correlation 0.52 with GFC, outperforming Correa (0.44) and Loughran-McDonald (0.46), comparable to Hubert (0.58)
- Cross-model LLM correlations of 0.90–0.93 for growth sentiment indicate robust consensus across architectures

Sentiment Validation: LLM Complements and Refines Lexical Methods

- LLMs complement established dictionaries by capturing structured, dimension-specific sentiment

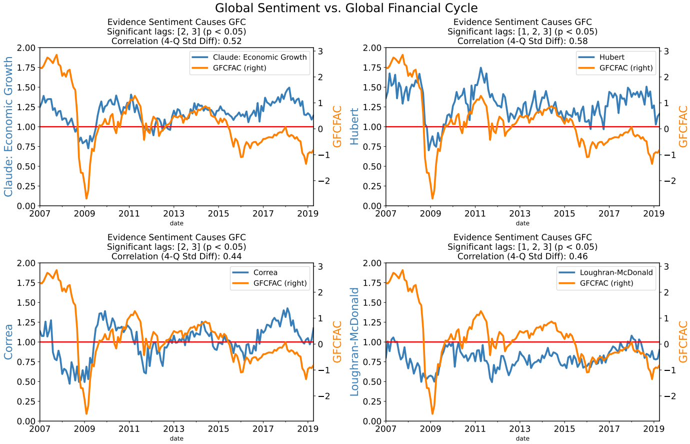
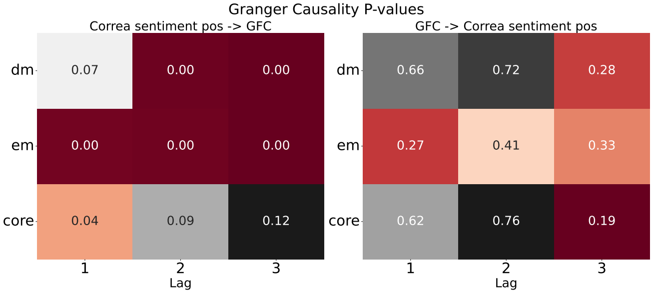


Figure 11: LLM-Generated Sentiment Indices vs. Lexical Approaches

Evidence that Sentiment Predicts the GFC

- Sentiment Granger-causes GFC ($p < 0.05$); GFC does *not* Granger-cause sentiment
- Granger causality captures **predictive precedence** rather than structural causation (Granger, 1969)
- Emerging-market CB sentiment appears to be a stronger predictor of the GFC than core CB sentiment
- Asymmetric pattern **consistent with** a communication transmission channel, though we cannot rule out omitted common factors or reverse causality at higher frequencies



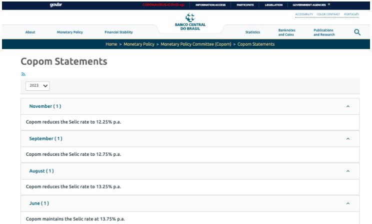
Outline

Summary & Future Work

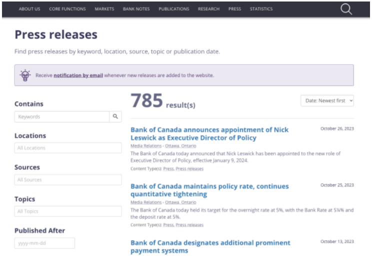
- We construct an open-source standardized database of CB communications designed as a benchmark for LLM tasks.
- We distribute all source code for our Data/LLM infrastructure to allow researchers to leverage AI tools in a scalable and reproducible manner.
- Using LLMs, we find that central banks often cite the U.S. economy broadly (rather than mention the U.S. Federal Reserve or monetary policy spillovers) as the primary driver of global conditions.
- Granger causality tests suggest that statement sentiment predicts the Global Financial Cycle rather than merely responding to it.
- We plan to scale our data to include more central banks and communication types.
- Our infrastructure includes major closed-source (Gemini, Claude, ChatGPT) and open-source (Hugging Face) LLM systems in an easy-to-use framework.

Outline

Monetary Policy Statements



(a) Brazil: Statements in press section



(b) Canada: Statements in press section

Figure 13: Examples of MPSs hosted on central bank websites

Key Words Time Series

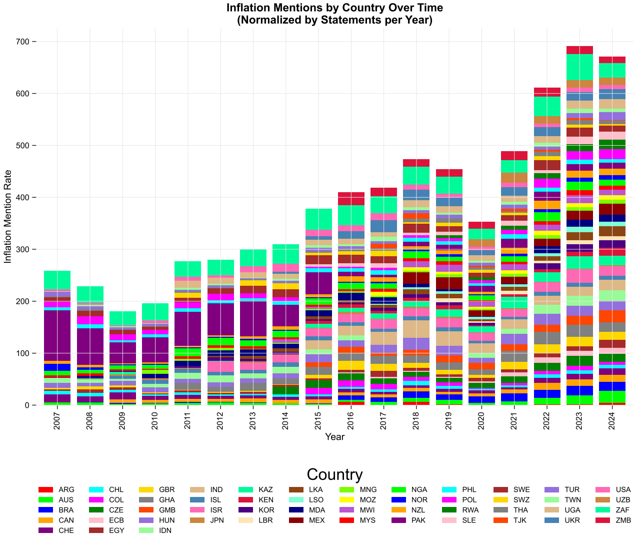


Figure 14: Inflation Mentions by Country

Regional Sentiment vs. Global Factor

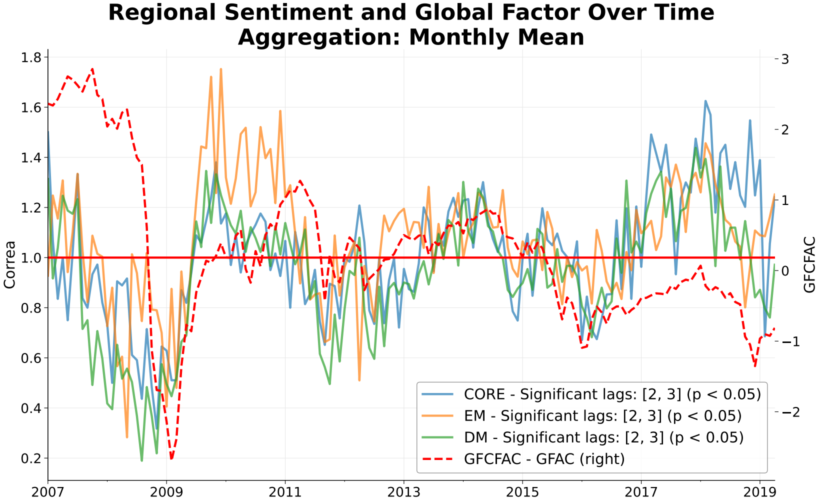


Figure 15: Regional Sentiment Evolution

References I

Born, B., Ehrmann, M., and Fratzscher, M. (2014). Central bank communication on financial stability. *Economic Journal*, 124(577):701–734.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–574.

Gonzalez, M. and Tadle, R. C. (2022). Monetary policy press releases: an international comparison. BIS Working Papers 1023, Bank for International Settlements.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Hansen, S. and McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99(S1):114–133.

References II

Miranda-Agrippino, S. and Rey, H. (2020). U.S. monetary policy and the Global Financial Cycle. *Review of Economic Studies*, 87(6):2754–2776.

Miranda-Agrippino, S. and Rey, H. (2022). The global financial cycle. In *Handbook of International Economics*, volume 6, pages 1–43. Elsevier.

Picault, M. and Renault, T. (2017). Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance*, 79:136–156.

Rey, H. (2015). Dilemma not trilemma: the global financial cycle and monetary policy independence. NBER Working Papers 21162, National Bureau of Economic Research, Inc.