

Words That Unite The World: A Unified Framework for Deciphering Central Bank Communications Globally

Agam Shah[✉], Siddhant Sukhani[✉], Huzaifa Pardawala[✉], Saketh Budideti[†], Riya Bhadani[†],
Rudra Gopal[†], Siddhartha Somani[†], Michael Galarnyk[†], Soungmin Lee[†],

Arnav Hiray, Akshar Ravichandran, Eric Kim, Pranav Aluru, Joshua Zhang,
Sebastian Jaskowski, Veer Guda, Meghaj Tarte, Liqin Ye, Spencer Gosden,
Rutwik Routu, Rachel Yuh, Sloka Chava, Sahasra Chava, Dylan Kelly,
Aiden Chiang, Harsit Mittal, Sudheer Chava

 Georgia Institute of Technology

✉ Corresponding and Equal First Authors: {ashah482, ssukhani3, hpardawala3}@gatech.edu

✉ and † Indicates core contributors.

 [Data and Models](#)  [Code](#)

Abstract

1 Central banks around the world play a crucial role in maintaining economic stability.
2 Deciphering policy implications in their communications is essential, especially
3 as misinterpretations can disproportionately impact vulnerable populations. To
4 address this, we introduce the World Central Banks (WCB) dataset, the most
5 comprehensive monetary policy corpus to date, comprising over 380k sentences
6 from 25 central banks across diverse geographic regions, spanning 28 years of
7 historical data. After uniformly sampling 1k sentences per bank (25k total) across
8 all available years, we annotate and review each sentence using dual annota-
9 tors, disagreement resolutions, and secondary expert reviews. We define three
10 tasks: Stance Detection, Temporal Classification, and Uncertainty
11 Estimation, with each sentence annotated for all three. We benchmark seven
12 Pretrained Language Models (PLMs) and nine Large Language Models (LLMs)
13 (Zero-Shot, Few-Shot, and with annotation guide) on these tasks, running 15,075
14 benchmarking experiments. We find that a model trained on aggregated data across
15 banks significantly surpasses a model trained on an individual bank’s data, con-
16 firming the principle “*the whole is greater than the sum of its parts.*” Additionally,
17 rigorous human evaluations, error analyses, and predictive tasks validate our frame-
18 work’s economic utility. Our artifacts are accessible through the HuggingFace, and
19 GitHub under the CC-BY-NC-SA 4.0 license.

20 1 Introduction

21 Over the past three decades, shocks such as the Dot-Com crash (2000–02), the Global Financial
22 Crisis (2008), the European Sovereign Debt Crisis (2010), negative oil prices (2020), the COVID-
23 19 pandemic (2020), and most recently, the Trump administration’s tariffs (2025) have left their
24 mark on the global economy. During these episodes, central banks responded by first stabilizing
25 volatile markets by expanding the money supply [Friedman, 2010, Buiter, 2008] and later by curbing
26 inflation by tightening monetary conditions [Hooley et al., 2023, Occhino, 2020]. The effectiveness

27 of these monetary policies hinges on central banks' ability to pursue a balance between two key goals:
 28 controlling inflation [Gerald, 2007] and maximizing employment [Epstein, 2007]. High inflation
 29 disproportionately harms lower-income populations, whose wages typically lag in catching up with
 30 rapidly rising prices. Conversely, inadequate employment severely limits economic opportunities for
 these same vulnerable groups, exacerbating social inequalities [Jayashankar and Murphy, 2023].

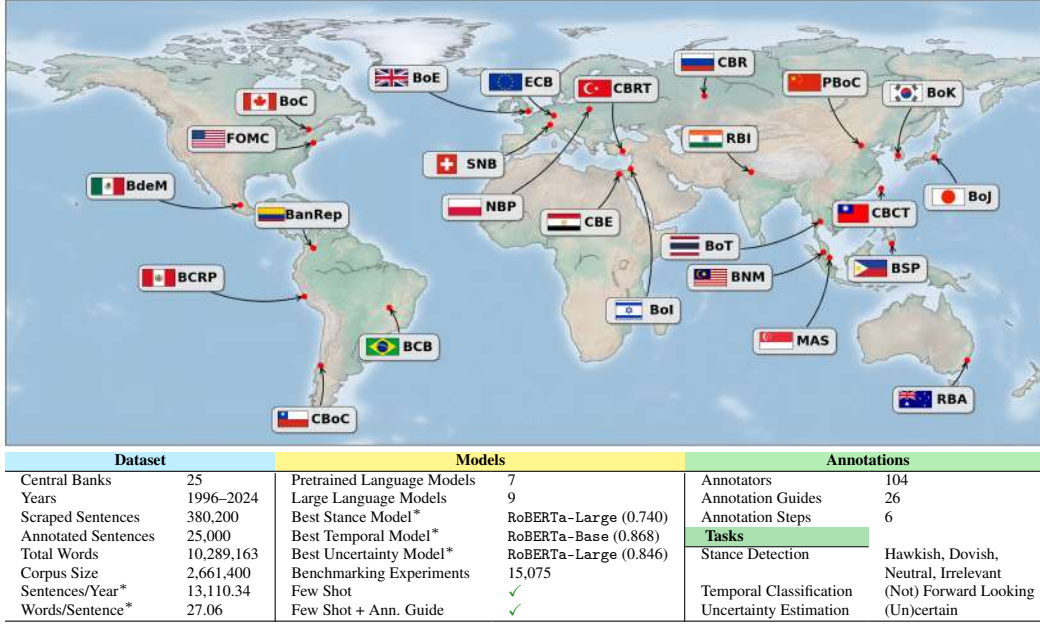


Figure 1: **A summary of the World Central Bank (WCB) dataset and experiments.** We systematically collect, clean and research the communications from 1996–2024 of 25 central banks at a sentence level, leading to 380,200 sentences (avg. 27.06 words/sentence) in our corpus. We present an annotated dataset consisting of 25,000 sentences across three tasks (Stance Detection, Temporal Classification, and Uncertainty Estimation) using comprehensive individual annotation guides and detailed instructions for annotation. We benchmark seven PLMs and eight LLMs on these tasks, under a bank-specific (1,000 bank-specific annotated sentences) and global setup (25,000 annotated sentences). The performance of the General (All-Banks) Setup model for each task is showcased in the figure. In these tables, * represents that it is an average.

31

32 However, despite the global implications of monetary policy, existing research and datasets used to
 33 train artificial intelligence (AI) models within the economic and financial space exhibit a pronounced
 34 geographical bias and substandard data quality [Tatarinov et al., 2025]. As highlighted by Longpre
 35 et al. [2024], most text datasets employed in financial and economic machine learning applications
 36 originate predominantly from web-crawled sources originating from a narrow geographical region,
 37 resulting in significant Western-centric biases. This limited representation within financial data
 38 restricts the generalizability and applicability of research insights globally [Manvi et al., 2024]. To
 39 address this critical gap, we introduce a new dataset WCB that significantly expands the scope of prior
 40 resources by collecting extensive monetary policy communications from a geographically diverse set
 41 of 25 central banks. Figure 1 shows a summary of our dataset and experiments. The dataset spans 28
 42 years and includes more than 380k sentences, averaging over 27 words each.

43 To ensure the utility and versatility of our dataset, we annotated 1k uniformly sampled sentences
 44 per bank (25k total) on three core tasks: Stance Detection, Temporal Classification,
 45 Uncertainty Estimation.. These labels jointly capture economic and monetary policy out-
 46 look (Stance Detection), when actions are signaled (Temporal Classification), and how
 47 confidently decisions are framed (Uncertainty Estimation), providing the essential dimensions
 48 for interpretation of central-bank communications across the world. Through these annotations and
 49 our experiments, our analysis reveals that the aggregated insights substantially exceed what could
 50 be inferred by considering each data subset independently. This “*whole is greater than the sum*
 51 *of its parts*” effect highlights the benefit of cross-bank transfer learning and confirms that training

on a unified corpus of monetary policy data captures shared linguistic patterns that individual bank models miss. Furthermore, we demonstrate the broad applicability of our best-performing fine-tuned PLMs beyond finance and economics, highlighting their potential for evaluating uncertainty and temporal orientation in other critical contexts, such as Congressional hearings on Conservation, Climate, Forestry, and Natural Resources (see Appendix D.4.2).

Our paper makes several key contributions to the existing literature. First, we introduce a comprehensive dataset of central bank communications alongside a systematic automated collection and cleaning method outlined in Section 3. This is coupled with extensive annotation guides specific to each bank (Appendix H) and thorough annotation instructions (Appendix B.6). Second, we present thorough annotation frameworks and release 78 fine-tuned PLMs capable of extracting nuanced insights from textual data, demonstrating their applicability across various contexts. We additionally perform extensive experiments and analysis under two evaluation setups: General (All-Banks) and Bank-Specific as defined in Section 4.2. These experiments include the generation of meeting minutes using LLMs (Appendix D.2), economic analysis with inflation data (Appendix E), error analysis (Appendix D.1) and performance gain analysis for Stance Detection model using the General Setup (Appendix C.6). Finally, our work highlights and addresses significant biases inherent in existing datasets and methodologies that focus on the specific central bank communications [Kumar et al., 2024, Tadde, 2022, Rozkrut et al., 2007, Kazinnik and Papell, 2021, Vega and Lahura, 2020, Oshima et al., 2018], as a result promoting greater transparency and responsible AI development. The subsequent sections detail our data collection and annotation processes, benchmarking frameworks, empirical results, and discussions of broader implications.

2 Task Definition

We define three sentence-level tasks designed to capture key dimensions of central bank communication: Stance Detection, Temporal Classification, and Uncertainty Estimation. The three tasks are defined as follows:

1. **Stance Detection** Each sentence is labeled as Hawkish, Dovish, Neutral (defined within Appendix G), or Irrelevant to capture the monetary stance of the central bank. The first three categories are adopted from Shah et al. [2023]. However, we introduce the Irrelevant class to capture salutations, logistical details, and operational jargon that do not convey any relevant information pertaining to monetary policy. This addition improves the quality of the collected corpus by preventing contamination from non-informative content without relying on a strict word dictionary. This is beneficial for our analysis as observed within Appendix B.9, wherein we illustrate the increased information retained by avoiding a word filter. Stance Detection serves as the central task in our work and is used across several experiments and analyses.
2. **Temporal Classification** Each sentence is classified as Forward Looking or Not Forward Looking. Forward Looking sentences refer to anticipatory and predictive sentences, often indicating future policy direction or economic expectations. Not Forward Looking sentences focus on past or contemporaneous observations. This task captures the temporal framing of the sentences and aids in understanding the temporal orientation of the sentence.
3. **Uncertainty Estimation** Each sentence is labeled as Certain or Uncertain, based on the presence of speculative language. It is designed to estimate the level of confidence conveyed in policy statements, which may reflect the clarity or ambiguity of the central bank’s economic outlook.

3 Dataset Construction

We develop our dataset using a thorough and systematic approach, adhering to strict methodological steps for both data collection and annotation, as shown in Figure 2 and detailed below.

Central Bank Selection We selected our initial set of banks using the central bank list provided by Sovereign Wealth Fund Institute [2025]. Filtering based on the banks’ assets in descending order, we prioritize central banks that had either monetary policy meeting minutes or equivalent documents in

both format and content available in English to ensure consistent annotations across central banks. A comprehensive reasoning on the selection of central banks is provided in Table 5 Appendix B.2. This selection process yielded a final corpus covering 26 central banks’ data, out of which one bank’s data (Czech National Bank) was chosen at random as a hold-out dataset for ablation studies. The data from the remaining 25 banks was used to create the actual dataset.

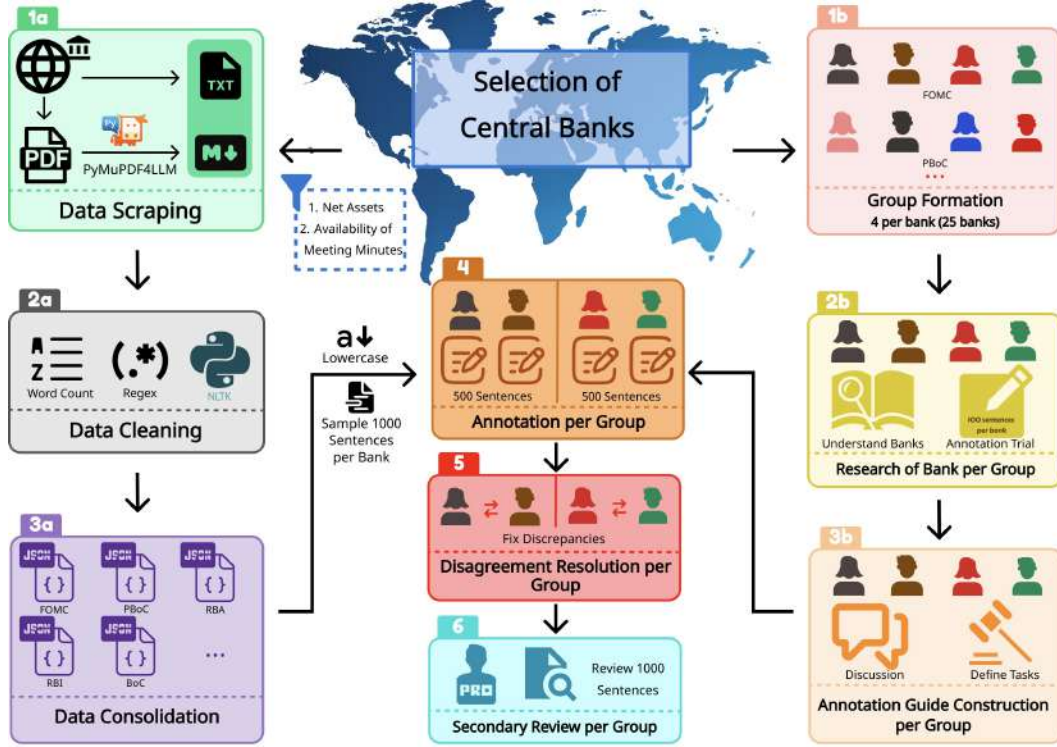


Figure 2: The dataset generation process for each central bank across the corpus of 25 central banks consists of three stages. (1a) We collected data from the central bank communications and converted them to Markdown files if they were PDFs. (2a) The data was cleaned using regex patterns and tokenized as described in Appendix B.3. (3a) The data for each central bank was consolidated into JSON files as described in B.5. (1b) Annotators were divided into 25 groups of four. (2b) They researched their assigned central bank and labeled a sample of 100 sentences. (3b) Individually drafted annotation guides by each group member were consolidated into a collaborative document for each bank. (4) Independently, the group members annotated the corresponding sentences using their annotation guide. (5) In pairs, they compared their annotations and resolved disagreements (Appendix H). (6) An expert annotator performed a final review of all the annotations by each group.

107

Data Collection, Cleaning, and Consolidation The raw documents were collected from the official central banks’ websites, targeting meeting minutes or their closest equivalents, as outlined in Table 6 Appendix B.3. Documents were extracted in either Markdown or plain text form and processed to retain only policy-relevant content. This involved removing contact details and other non-textual artifacts. The cleaned sentences from these documents were then consolidated into JSON files, with full metadata for each document stored as outlined in Appendix B.5. The full collection, cleaning, and consolidation pipeline is described extensively in Appendix B.3.

Annotation Methodology Whilst the data was being preprocessed, the entire team of annotators was split into subgroups (four annotators per bank) and assigned various responsibilities required to complete the annotation procedure. Instructions given to the annotators and the full sequence of steps leading to the final annotations are outlined in Appendix B.6, including participation terms, initial research, annotation guide development, annotation steps, and disagreement resolution. A key contribution of the dataset also lies in the creation a thorough annotation guide frameworks, to the

best of our knowledge. For each of the 26 central banks, annotators developed bank-specific labeling criteria after conducting extensive research on monetary policy mandates, economic history, and communication style. Initial annotation guides were drafted independently and then consolidated into a detailed document. Annotations were carried out in a double-blind manner, with pairwise comparison for disagreement resolution. This was followed by the review of all annotations by an expert. This entire process is visualized in Figure 2. Further details on the annotation guides, annotator names and contact information, as well as inter-annotator agreement statistics for all three tasks are provided in Appendix H. These details are reported separately for each central bank.

4 Benchmarking

To evaluate the performance of Language Models (LMs) on the tasks (as defined in Section 2), we benchmark a suite of seven PLMs and nine LLMs on our annotated dataset, with full specifications provided in Table 11 Appendix C.2. For each experiment, we use a 700-150-150 train-test-validation split. For PLMs, we run a grid search across the following parameters: seeds = [5768, 78516, 944601]¹, batch sizes = [32, 16], learning rates = [1e-5, 1e-6]. Full training specifications, compute environment, and LLM API usage are detailed in Appendix C.3, and the different prompts used for LLM inference are listed in Appendix C.1.

4.1 Models

The list of models used is given below. Model-specific details are provided in Table 11, Appendix C.2.

PLMs Bert-Large [Devlin et al., 2019], BERT-Base [Devlin et al., 2019], RoBERTa-Large [Liu et al., 2019], RoBERTa-Base [Liu et al., 2019], ModernBert-large [Warner et al., 2024], ModernBert-Base [Warner et al., 2024], FinBert-Pretrain [Huang et al., 2023].

Closed Source LLMs Gemini 2.0 flash [Google DeepMind, 2024], GPT-4o (gpt-4o-2024-08-06) [Hurst et al., 2024], GPT 4.1 (gpt-4.1-2025-04-14) and GPT 4.1 mini (gpt-4.1-mini-2025-04-14) [OpenAI, 2025].

Open Source (open weights) LLMs Llama-3-70B-Chat [Dubey et al., 2024], DeepSeek-V3 (DeepSeek-V3-0324) [Liu et al., 2024], Qwen2.5-72B-Instruct-Turbo [Bai et al., 2023], FinMA [Xie et al., 2023] and Llama-4-Scout-17B-16E-Instruct [AI, 2024a].

Domain Specific Models FinBert-Pretrain [Huang et al., 2023] and FinMA [Xie et al., 2023].

4.2 Results

We present benchmarking results for all three tasks: Stance Detection, Temporal Classification, and Uncertainty Estimation across multiple PLMs and LLMs. LLMs were evaluated under one setup whereas PLMs were evaluated under two distinct setups defined as follows:

General (All-Banks) Setup In this setup, models are evaluated on a single combined dataset containing all annotated sentences from all 25 central banks. As a preliminary experiment, we vary training set size and observe that the F1-Scores plateau for Stance Detection when the set size is 700 (Appendix C.7).

The benchmarking results for Stance Detection are shown in Table 1, Temporal Classification in Table 12 Appendix C.4.1, and Uncertainty Estimation in Table 13 Appendix C.4.1.

Bank-Specific Setup In this setup, each central bank’s data is treated as an independent dataset. Models are evaluated separately on each bank’s data, and the average weighted F1-Score across all 25 banks is reported. Results for Stance Detection are shown in Table 14, Temporal Classification in Table 15 Appendix C.4.2, and Uncertainty Estimation in Table 16 Appendix C.4.2.

¹These specific seeds were chosen based on the work of Shah et al. [2022] and Shah et al. [2023].

Table 1: F1-Scores across the banks for the Stance Detection label of Hawkish, Dovish, Neutral, Irrelevant with the standard deviation in brackets. These results are based on the All-Banks General Dataset Setup, as mentioned in Section 4.2, on each bank. The best performing PLM and LLM in each row is highlighted in blue and green respectively. The best performing model for a given bank is bold. The model names are abbreviated in accordance with Table 11, Appendix C.2. The LLMs are evaluated using a zero-shot prompt.

Bank	Base				Large			Closed-Source				Open-Source				
	MBB	BB	FB	RBB	MBL	BL	RBL	Gem	4o	4.1m	4.1	DS	Qwen	FM	L3	L4S
BCB	.678 (.039)	.635 (.057)	.609 (.008)	.655 (.018)	.673 (.016)	.636 (.014)	.634 (.050)	.528 (.027)	.498 (.017)	.462 (.020)	.504 (.017)	.613 (.021)	.525 (.028)	.350 (.034)	.503 (.021)	.589 (.049)
BCRP	.788 (.021)	.781 (.008)	.764 (.009)	.779 (.028)	.798 (.013)	.801 (.024)	.821 (.015)	.675 (.004)	.628 (.008)	.634 (.035)	.665 (.004)	.666 (.010)	.503 (.062)	.301 (.031)	.641 (.014)	.620 (.043)
BNM	.626 (.029)	.629 (.017)	.653 (.023)	.601 (.012)	.630 (.041)	.644 (.013)	.640 (.009)	.409 (.006)	.443 (.027)	.430 (.007)	.409 (.025)	.475 (.013)	.333 (.026)	.160 (.033)	.567 (.025)	.435 (.005)
BSP	.741 (.015)	.697 (.020)	.707 (.028)	.749 (.010)	.724 (.049)	.698 (.025)	.784 (.017)	.424 (.042)	.420 (.069)	.451 (.039)	.514 (.076)	.534 (.029)	.380 (.015)	.245 (.027)	.584 (.042)	.500 (.035)
BanRep	.685 (.031)	.638 (.022)	.679 (.017)	.673 (.013)	.702 (.001)	.650 (.017)	.701 (.013)	.515 (.021)	.455 (.031)	.520 (.036)	.553 (.015)	.570 (.037)	.450 (.023)	.230 (.033)	.573 (.038)	.423 (.033)
BoC	.722 (.033)	.740 (.006)	.750 (.010)	.745 (.032)	.721 (.004)	.755 (.005)	.785 (.010)	.629 (.069)	.647 (.052)	.641 (.052)	.657 (.028)	.657 (.026)	.524 (.038)	.264 (.029)	.669 (.043)	.644 (.024)
BoE	.686 (.031)	.706 (.056)	.734 (.052)	.769 (.019)	.735 (.039)	.755 (.009)	.785 (.034)	.543 (.026)	.524 (.031)	.537 (.048)	.602 (.031)	.543 (.070)	.396 (.021)	.129 (.026)	.661 (.044)	.518 (.045)
Bol	.652 (.003)	.642 (.023)	.616 (.019)	.614 (.008)	.658 (.023)	.628 (.033)	.689 (.017)	.474 (.011)	.460 (.005)	.433 (.032)	.526 (.013)	.482 (.001)	.329 (.025)	.085 (.011)	.594 (.023)	.430 (.008)
BoJ	.691 (.020)	.662 (.044)	.629 (.047)	.660 (.020)	.708 (.042)	.683 (.033)	.702 (.025)	.524 (.010)	.545 (.021)	.465 (.028)	.565 (.008)	.498 (.009)	.406 (.040)	.157 (.033)	.574 (.027)	.507 (.026)
BoK	.723 (.056)	.664 (.009)	.679 (.011)	.706 (.018)	.740 (.009)	.700 (.028)	.755 (.019)	.646 (.040)	.648 (.030)	.594 (.066)	.678 (.016)	.629 (.016)	.466 (.076)	.181 (.031)	.632 (.032)	.592 (.047)
BdeM	.696 (.020)	.694 (.048)	.642 (.032)	.724 (.023)	.716 (.008)	.684 (.026)	.735 (.030)	.596 (.009)	.602 (.024)	.509 (.023)	.626 (.016)	.669 (.034)	.447 (.047)	.118 (.018)	.642 (.012)	.552 (.013)
BoT	.717 (.046)	.696 (.064)	.723 (.021)	.735 (.057)	.733 (.071)	.728 (.006)	.741 (.017)	.547 (.004)	.549 (.032)	.551 (.039)	.573 (.012)	.581 (.029)	.484 (.038)	.258 (.026)	.596 (.009)	.577 (.030)
CBCT	.641 (.049)	.637 (.031)	.635 (.029)	.667 (.015)	.616 (.035)	.678 (.032)	.688 (.024)	.451 (.044)	.474 (.044)	.474 (.031)	.485 (.026)	.522 (.022)	.388 (.049)	.180 (.037)	.556 (.015)	.475 (.051)
CBE	.773 (.027)	.783 (.026)	.790 (.004)	.810 (.016)	.822 (.015)	.788 (.029)	.836 (.014)	.629 (.037)	.672 (.036)	.581 (.045)	.648 (.036)	.636 (.007)	.352 (.056)	.142 (.014)	.702 (.021)	.594 (.024)
CBR	.763 (.031)	.754 (.017)	.750 (.020)	.798 (.032)	.811 (.023)	.779 (.022)	.835 (.011)	.759 (.015)	.749 (.027)	.693 (.026)	.701 (.049)	.794 (.028)	.573 (.035)	.146 (.022)	.772 (.029)	.665 (.013)
CBRT	.717 (.015)	.743 (.016)	.724 (.012)	.746 (.011)	.746 (.014)	.724 (.032)	.762 (.027)	.495 (.006)	.421 (.014)	.424 (.018)	.475 (.015)	.539 (.030)	.277 (.006)	.133 (.020)	.653 (.036)	.416 (.032)
CBoC	.760 (.048)	.747 (.006)	.743 (.054)	.779 (.049)	.792 (.043)	.793 (.058)	.799 (.037)	.668 (.032)	.604 (.027)	.605 (.033)	.678 (.057)	.676 (.038)	.559 (.072)	.223 (.040)	.685 (.019)	.539 (.071)
ECB	.707 (.040)	.699 (.048)	.668 (.030)	.724 (.050)	.724 (.014)	.713 (.022)	.755 (.024)	.638 (.023)	.599 (.019)	.610 (.038)	.660 (.021)	.637 (.005)	.548 (.017)	.206 (.052)	.613 (.020)	.595 (.016)
FOMC	.671 (.029)	.674 (.035)	.675 (.046)	.747 (.020)	.732 (.037)	.685 (.044)	.749 (.047)	.572 (.021)	.584 (.025)	.564 (.018)	.649 (.023)	.653 (.023)	.512 (.023)	.170 (.025)	.599 (.012)	.498 (.015)
MAS	.656 (.049)	.681 (.043)	.666 (.005)	.666 (.066)	.698 (.042)	.680 (.049)	.703 (.033)	.553 (.046)	.581 (.026)	.588 (.034)	.569 (.015)	.689 (.041)	.540 (.026)	.347 (.034)	.638 (.035)	.646 (.023)
NBP	.685 (.016)	.690 (.013)	.705 (.026)	.731 (.013)	.725 (.009)	.697 (.010)	.695 (.020)	.637 (.015)	.631 (.043)	.614 (.063)	.665 (.031)	.660 (.002)	.508 (.028)	.118 (.017)	.618 (.035)	.597 (.015)
PBoC	.791 (.033)	.763 (.008)	.742 (.046)	.787 (.026)	.813 (.014)	.793 (.007)	.786 (.017)	.492 (.046)	.559 (.037)	.531 (.026)	.535 (.033)	.592 (.037)	.379 (.017)	.128 (.018)	.613 (.033)	.446 (.024)
RBA	.685 (.023)	.681 (.019)	.682 (.027)	.672 (.015)	.707 (.029)	.695 (.024)	.741 (.028)	.531 (.049)	.478 (.079)	.483 (.058)	.553 (.074)	.537 (.055)	.358 (.049)	.133 (.020)	.614 (.034)	.495 (.057)
RBI	.604 (.037)	.649 (.035)	.653 (.041)	.628 (.034)	.633 (.032)	.655 (.039)	.668 (.044)	.489 (.025)	.519 (.041)	.509 (.026)	.495 (.027)	.542 (.016)	.431 (.008)	.231 (.058)	.581 (.030)	.557 (.032)
SNB	.692 (.017)	.685 (.028)	.653 (.016)	.685 (.009)	.725 (.018)	.698 (.030)	.713 (.020)	.635 (.003)	.601 (.015)	.640 (.037)	.652 (.016)	.643 (.024)	.554 (.029)	.252 (.008)	.612 (.014)	.607 (.051)
Average	.702 (.030)	.695 (.028)	.691 (.025)	.714 (.025)	.723 (.026)	.710 (.025)	.740 (.024)	.562 (.085)	.556 (.084)	.542 (.075)	.586 (.078)	.601 (.075)	.449 (.083)	.196 (.071)	.620 (.053)	.541 (.074)

4.3 Model Performance Analysis

PLM Performance Analysis For PLMs in the General Setup for Stance Detection, we observe superior performance from RoBERTa-Large, achieving an average weighted F1-Score of 0.740. Moreover, larger models consistently outperform their base model counterparts.

Interestingly, ModernBERT-Large does not emerge as the best-performing model for most banks. It only achieves the highest performance for four banks. This suggests that, for specific domains, it may not generalize well, and RoBERTa models (Base and Large) offer greater performance for this particular task. Another significant observation is the relatively poor performance of FinBERT-Pretrain, despite being fine-tuned on financial data. Its lower than average weighted F1-Score indicates that our Stance Detection task involves complexities beyond the recognition of financial jargon, highlighting substantial differences from traditional sentiment analysis tasks and emphasizing the nuanced linguistic complexities inherent to monetary policy communications.

LLM Performance Analysis (Zero-Shot prompting) Using Zero-Shot prompting, Llama-3-70B outperforms all other LLMs, achieving an average weighted F1-Score of 0.620 on the Stance Detection task. In contrast, Llama-4-Scout performs worse across all banks, showing limited domain adaptation in this context. GPT-4.1 and DeepSeek-V3 perform the best for a combined total of eight banks, while Gemini-2.0-Flash achieves top performance on one bank. The domain-specific model FinMA has the lowest average performance across all LLMs. We attribute this to the absence of monetary policy-specific data in FinMA’s training corpus. For Stance Detection, Table 1 clearly shows that PLMs consistently outperform LLMs using Zero-Shot prompting on average.

5 Additional Experiments

Performance Gain Analysis: The whole is greater than the sum of its parts To understand why the best-performing General Setup model consistently outperforms the best-performing Bank-Specific Setup models for the Stance Detection task, we conduct qualitative and statistical analyses. For the qualitative analysis, we focus on test-set sentences that were misclassified by the best Bank-Specific model but correctly classified by the General model (RoBERTa-Large). For each bank, we sample such sentences, compute their sentence embeddings using voyage-finance-2 model AI [2024b], and find the most semantically similar sentence in the training corpus. Representative examples within Table 23 Appendix C.6.1 show that the General Setup model benefits from semantically related sentences across banks, allowing it to learn from a much larger and diverse sample.

For the statistical analysis, we compute TF-IDF vectors for each document and calculate cross-bank cosine similarity at a document level. Regressing the change in weighted F1-Score (General Setup – Specific Setup) for RoBERTa-Large on normalized average similarity reveals a statistically significant relationship ($p = 0.016$), as shown in Appendix C.6.2. This confirms that semantic similarity across banks correlates with larger gains from the General Setup model. Together, these results (discussed in Appendix C.6) confirm that the General Setup model leverages cross-bank semantic patterns to improve performance. This implies that the more a bank’s language resembles that of others, the more it benefits from a model trained on the aggregated dataset. This highlights the value of shared linguistic structure in monetary policy communication.

Error Analysis We conduct error analysis on the annotated sentences from People’s Bank of China, focusing on the the Zero-Shot output from the best performing LLM (Llama-3-70B-Chat) on the three tasks. We select this central bank for two reasons: (1) it represents one of the world’s largest banks in terms of assets [Sovereign Wealth Fund Institute, 2025] and (2) we find that while the overall policy tone is generally dovish, many statements appear linguistically neutral, making Stance Detection particularly challenging. We examine the subset of sentences that were misclassified by Llama-3-70B-Chat, aiming to characterize common errors in judgment. The split of selected sentences, types of errors, analysis procedure, and detailed results are provided in Appendix D.1.

Benchmarking with Few-Shot and prompting with Annotation Guide In Section 4.3, LLMs are evaluated using basic Zero-Shot prompting. To analyze the impact of sophisticated prompting strategies on LLM performance across all three tasks, we further evaluate the best-performing LLM on average (Llama-3-70B-Chat) using two advanced prompting strategies: Few-Shot prompting and prompting with Annotation Guide. The detailed procedures and prompt templates for Few-Shot prompting are provided in C.1.4 and for prompting with Annotation Guide are provided in Appendix C.1.5. As reported in Table 17 Appendix C.4, both prompting strategies result in marginal improvements on the Stance Detection task but lead to performance declines on Temporal Classification and Uncertainty Estimation. A detailed breakdown of results is provided in Appendix C.4.

Generating Meeting Minutes Documents Using an LLM To extend our analysis beyond the annotated dataset, we generate meeting minutes for a target document whose release date is beyond the knowledge cutoff of May 31st, 2024. We use GPT-4.1, selected for its one million token context window and strong performance relative to other models with sufficiently large context windows on the Stance Detection task. Each generation includes example inputs, an example output and actual context inputs as defined in Appendix D.2.1; due to the frequency and timing of meetings which may differ across various central banks, some documents may overlap across the example

and actual input sets. We evaluate the generated minutes by computing their hawkishness scores, as defined in Section 6, and compare them to the hawkishness measures of the corresponding target documents labeled by our best-performing fine-tuned Stance Detection model. This quantifies alignment with real-world policy tone and demonstrates a novel use case for our fine-tuned models: as evaluators of generated policy content. The full generation procedure is discussed in D.2.1 and associated results are presented in Appendix D.2.2.

Human Evaluation To benchmark against human performance, we conduct an evaluation on the European Central Bank’s (ECB) communications, selected due to its direct influence on 20 member nations. We compare the best-performing LLM for each task: Llama-3-70B-Chat for Temporal Classification and Uncertainty Estimation and GPT 4.1 for Stance Detection against a human annotator without access to the annotation guide. Using 50 samples per task, the LLM consistently outperforms the human across all three tasks, despite being evaluated with Zero-Shot prompting. The methodology for this experiment as well as results are detailed in Appendix D.3.

6 Economic Analysis

To understand the impact that the stance of a central bank’s communications can have on an economy, we inference our best performing model for Stance Detection (RoBERTa-Large) on our corpus of over 380k sentences and generate the following measure of hawkishness Shah et al. [2023] at a document level and compare it to the inflation within the specific central bank’s economy.

$$\text{Hawkishness Measure} = \frac{\# \text{ Hawkish Sentences} - \# \text{ Dovish Sentences}}{\# \text{ Total Sentences} - \# \text{ Irrelevant Sentences}} \quad (1)$$

This was computed for each document associated with every central bank. As seen in Appendix E, we then plot the change of hawkishness measure over time alongside the percentage change in Consumer Price Index (CPI) for each economy to determine if our measure aligns with individual economies’ inflation. Building on the findings of Shah et al. [2023], Kumar et al. [2024], our analysis shows that the hawkishness measures derived from monetary policy documents closely track inflation trends within each economy, particularly during periods of financial distress such as the 2001 and 2008 crises and the COVID-19 pandemic as seen in Appendix E.

7 Related Works

Table 2: Comparison of related datasets. The *Corpus* is the aggregated data size and the number of DP (*Data Points*) is the annotated dataset size. CB (*Central Banks*) is the number of central banks included in the data set. *Data* is the content which was analyzed in the paper. S (*Stance*), T (*Temporal*), U (*Uncertainty*) specify if the corresponding label is present. AG (*Annotation Guide*) details if an annotation guide was used and is provided.

Paper	Year Range	Corpus	DP	CB	Data	S	T	U	AG
Ivrendi and Yildirim [2013]	1995–2012	2,544	2,544	6	Policy Transcripts	×	×	×	×
Armeliu et al. [2020]	2002–2017	12,024	–	23	Speeches, Transcripts	✓	×	×	×
Mathur et al. [2022]	2009–2022	24,180	24,180	6	MP Conferences	✓	×	×	×
Kirti et al. [2022]	2020	5,462	5,462	74	Gov. Websites, MP Reports, News	×	×	×	×
Shah et al. [2023]	1996–2022	216,926	7,440	21	FOMC Minutes, Speeches, Conferences	✓	×	×	✓
Bolhuis [2024]	2000–2022	3,545	3,545	20	MP Announcements, Finance Data	×	×	×	×
Barigozzi et al. [2024]	2000–2024	800	800	1	National Accounts	×	×	×	×
Zhang et al. [2025]	2008–2024	25,690	–	0	MacroEcon Announcements, Finance data	×	×	×	×
WCB (Ours)	1996–2024	2,661,400	100,000	25	Monetary Policy Minutes	✓	✓	✓	✓

Over the last two decades, Natural Language Processing (NLP) has significantly impacted fields like finance and economics [Tatarinov et al., 2025, Ehrmann and Fratzscher, 2007]. In finance, developments such as specialized financial sentiment lexicons [Loughran and McDonald, 2010] have enabled detailed analyses of textual datasets, facilitating research on monetary policy, Uncertainty Estimation [Gentzkow et al., 2019, Baker et al., 2016] and Temporal Classification [Hansen et al., 2019], and monetary policy Stance Detection [Shah et al., 2023].

Central bank communications, including press conferences, speeches, minutes, and economic news, have been analyzed using NLP [Gorodnichenko et al., 2023a, Barbaglia et al., 2022, Ahrens et al.,

263 2024, Apel and Grimaldi, 2014, Ehrmann and Fratzscher, 2007, Shah et al., 2023, Pardawala et al.,
 264 2025]. These studies help identify monetary policy shocks [Aruoba and Drechsel, 2024], assess
 265 sentiment [Hilscher et al., 2024], and explore the implications of transparency on market expectations
 266 [Gorodnichenko et al., 2023a]. Recent works leverage transformer-based models like BERT and
 267 RoBERTa [Shah et al., 2023, Gambacorta et al., 2024] alongside methods like LDA [Jegadeesh and
 268 Wu, 2017], word embeddings [Matsui et al., 2021], and graph clustering [Zhou et al., 2017].

269 Current analyses often focus on individual central banks such as RBI, FOMC, and ECB, providing
 270 localized insights [Kumar et al., 2024, Tadle, 2022, Tobback et al., 2017, Kazinnik and Papell, 2021,
 271 Vega and Lahura, 2020, Rozkrut et al., 2007, Oshima et al., 2018, Iglesias et al., 2017, Aguilar and
 272 Pérez-Cervantes, 2022]. This approach, however, limits the understanding of broader global trends
 273 [Ehrmann and Talmi, 2020, Wong and Fong, 2011, Armelius et al., 2020]. Our study fills this gap by
 274 analyzing texts from 25 central banks across multiple economic crises, offering comprehensive global
 275 insights. Table 2 shows that our dataset represents the largest monetary policy dataset created to date.
 276 Additional related work and detailed dataset comparisons are available in Table 30 Appendix F.

277 8 Discussion

278 To the best of our knowledge, this is the first dataset in the finance and economics domain covering
 279 26 central banks across six continents with 25k annotated sentences (more than 380k sentences
 280 total), structured metadata (Appendix B.5) and three core tasks. We provide detailed, bank-specific
 281 annotation guides (Appendix H), annotator instructions (Appendix B.6), and documentation of
 282 the annotation process (Section 3) to facilitate reproducibility. Beyond being a benchmark, our
 283 work enables rich downstream analysis: improved performance via generalization across banks
 284 (Appendix C.6), meeting minutes’ generation using LLMs (Appendix D.2), economic analysis
 285 (Appendix E), human evaluation (Appendix D.3), and error analysis (Appendix D.1). We also show
 286 that fine-tuned PLMs outperform LLMs on all three tasks, even with Few-Shot or Annotation-Guide
 287 prompting (Appendix C.5).

288 **Global Coverage Gaps** Despite our efforts to ensure wide geographical representation, the broader
 289 data ecosystem remains skewed. As highlighted by Longpre et al. [2024], over 93% of textual data
 290 (tokens) still originates from North American and European institutions, while data from Africa and
 291 South America accounts for less than 0.2% of content. This persistent imbalance limits the robustness
 292 and fairness of language technologies. Our dataset takes a step toward bridging this gap, but more
 293 work is needed to balance representation.

294 **Broader Impacts** To demonstrate broader applicability of our work **beyond finance**, we evaluate
 295 the Temporal Classification and Uncertainty Estimation tasks on U.S. Congressional
 296 Committee hearings. Shown in Appendix D.4.2, our best-performing PLM achieves F1-Scores of
 297 0.879 and 0.683, respectively, indicating strong generalization. Additionally, to assess transferability
 298 **beyond the chosen 25 central banks’ data**, we evaluate the Stance Detection task on cleaned
 299 data from the 26th bank, Czech National Bank, and observe a mean weighted F1-Score of 0.800,
 300 as detailed in Appendix D.4.1. This confirms that our models retain high performance even on
 301 out-of-sample bank’s data.

302 **Open Access and Expansion** We currently cover 26 central banks and initiate this project along
 303 with its publicly available Hugging Face repository and GitHub repository as an open research effort.
 304 Our goal is to expand beyond the current 26 central banks through community collaboration and
 305 lower entry barriers, especially for independent researchers. To support this, we also provide cleaned
 306 data for two additional central banks (Appendix B.2) along with a reproducible pipeline that enables
 307 others to build upon and extend our work.

308 Acknowledgments and Disclosure of funding

309 We would like to thank all the annotators who annotated the sentences and contributed to initial drafts
 310 of the annotation guides. We also appreciate the Georgia Tech Library for giving us access to the
 311 Bloomberg Terminal, Partnership for an Advanced Computing Environment (PACE) at Georgia Tech,
 312 and especially Label Studio.

References

- Ana Aguilar and Fernando Pérez-Cervantes. Communication, monetary policy, and financial markets in Mexico. Technical report, Bank for International Settlements, 2022.
- Maximilian Ahrens, Deniz Erdemlioglu, Michael McMahon, Christopher J. Neely, and Xiye Yang. Mind your language: Market responses to central bank speeches. *Journal of Econometrics*, page 105921, 2024. ISSN 0304-4076. doi: <https://doi.org/10.1016/j.jeconom.2024.105921>. URL <https://www.sciencedirect.com/science/article/pii/S0304407624002720>.
- Meta AI. Llama 4 scout 17b 16e instruct. <https://huggingface.co/meta-llama/Llama-4-Scout-17B-16E-Instruct>, 2024a. Accessed: 2025-04-12.
- Voyage AI. voyage-finance-2: Domain-specific embedding model for finance. <https://blog.voyageai.com/2024/06/03/domain-specific-embeddings-finance-edition-voyage-finance-2/>, 2024b. Accessed: 2025-05-12.
- Mikael Apel and Marianna Blix Grimaldi. How informative are central bank minutes? *Review of Economics*, 65(1):53–76, 2014. doi: [doi:10.1515/roe-2014-0104](https://doi.org/10.1515/roe-2014-0104). URL <https://doi.org/10.1515/roe-2014-0104>.
- Hanna Armelius, Christoph Bertsch, Isaiah Hull, and Xin Zhang. Spread the word: International spillovers from central bank communication. *Journal of International Money and Finance*, 103:102116, 2020. ISSN 0261-5606. doi: <https://doi.org/10.1016/j.jimonfin.2019.102116>. URL <https://www.sciencedirect.com/science/article/pii/S0261560619302967>.
- S Borağan Aruoba and Thomas Drechsel. Identifying monetary policy shocks: A natural language approach. Technical report, National Bureau of Economic Research, 2024.
- Jinze Bai, Shuai Bai, Yunfei Chu, Zeyu Cui, Kai Dang, Xiaodong Deng, Yang Fan, Wenbin Ge, Yu Han, Fei Huang, Binyuan Hui, Luo Ji, Mei Li, Junyang Lin, Runji Lin, Dayiheng Liu, Gao Liu, Chengqiang Lu, Keming Lu, Jianxin Ma, Rui Men, Xingzhang Ren, Xuancheng Ren, Chuanqi Tan, Sinan Tan, Jianhong Tu, Peng Wang, Shijie Wang, Wei Wang, Shengguang Wu, Benfeng Xu, Jin Xu, An Yang, Hao Yang, Jian Yang, Shusheng Yang, Yang Yao, Bowen Yu, Hongyi Yuan, Zheng Yuan, Jianwei Zhang, Xingxuan Zhang, Yichang Zhang, Zhenru Zhang, Chang Zhou, Jingren Zhou, Xiaohuan Zhou, and Tianhang Zhu. Qwen technical report, 2023. URL <https://arxiv.org/abs/2309.16609>.
- Scott R Baker, Nicholas Bloom, and Steven J Davis. Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636, 2016.
- Caroline Banton, Amy Drury, and Yariet Perez. Interest rates: Types and what they mean to borrowers. <https://www.investopedia.com/terms/i/interestrate.asp>, 2025. Accessed 7-May-2025.
- Luca Barbaglia, Sergio Consoli, and Sebastiano Manzan. Forecasting with economic news, 2022. URL <https://arxiv.org/abs/2203.15686>.
- Matteo Barigozzi, Claudio Lissone, and Lorenzo Tonni. Large datasets for the euro area and its member countries and the dynamic effects of the common monetary policy. *arXiv preprint arXiv:2410.05082*, 2024.
- BerriAI. LiteLLM: Python sdk and proxy server for unified access to 100+ llm apis. <https://github.com/BerriAI/litellm>, 2024. Version 1.68.0, accessed 5-May-2025.
- Marijn Bolhuis. A New Dataset of High-Frequency Monetary Policy Shocks. *IMF Working Papers*, 2024(224): 1, 10 2024. ISSN 1018-5941. doi: [10.5089/9798400290930.001](https://dx.doi.org/10.5089/9798400290930.001). URL <http://dx.doi.org/10.5089/9798400290930.001>.
- Willem H Buiter. Central banks and financial crises. 2008.
- Prafulla Choubey, Alexander Fabbri, Caiming Xiong, and Chien-Sheng Wu. Lexical repetitions lead to rote learning: Unveiling the impact of lexical overlap in train and test reference summaries. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 7269–7283, Singapore, December 2023. Association for Computational Linguistics. doi: [10.18653/v1/2023.findings-emnlp.484](https://doi.org/10.18653/v1/2023.findings-emnlp.484). URL <https://aclanthology.org/2023.findings-emnlp.484/>.
- Olivier Coibion, Yuriy Gorodnichenko, and Michael Weber. Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, 130(6):1537–1584, 2022. doi: [10.1086/718982](https://doi.org/10.1086/718982). URL <https://doi.org/10.1086/718982>.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024.

Michael Ehrmann and Marcel Fratzscher. Communication by central bank committee members: different strategies, same effectiveness? *Journal of Money, Credit and Banking*, 39(2-3):509–541, 2007.

Michael Ehrmann and Jonathan Talmi. Starting from a blank page? semantic similarity in central bank communication and market volatility. *Journal of Monetary Economics*, 111:48–62, 2020.

Gerald Epstein. Central banks as agents of employment creation. 2007.

European Central Bank. Our money, 2025. URL <https://www.ecb.europa.eu/euro/intro/html/index.en.html>. Accessed: 2025-04-03.

Eurostat. Hicp methodological manual — 2024 edition. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/w/ks-gq-24-003>, 2024. Accessed 7-May-2025.

Jason Fernando, Michael J Boyle, and Pete Rathburn. Inflation: What it is, how it works, and how it’s measured. <https://www.investopedia.com/terms/i/inflation.asp>, 2024. Accessed 7-May-2025.

Jason Fernando, Peter Westfall, and Pete Rathburn. Consumer price index (cpi): What it is and how it’s used. <https://www.investopedia.com/terms/c/consumerpriceindex.asp>, 2025. Accessed 7-May-2025.

Benjamin Friedman. Learning from the crisis: What can central banks do? 2010.

Leonardo Gambacorta, Byeungchun Kwon, Taejin Park, Pietro Patelli, and Sonya Zhu. *CB-LMs: language models for central banking*. Bank for International Settlements, Monetary and Economic Department, 2024.

Matthew Gentzkow, Bryan Kelly, and Matt Taddy. Text as data. *Journal of Economic Literature*, 57(3):535–74, September 2019. doi: 10.1257/jel.20181020. URL <https://www.aeaweb.org/articles?id=10.1257/jel.20181020>.

Epstein Gerald. Central bank, inflation targeting and employment creation, 2007.

Google DeepMind. Gemini 2.0 flash, 2024. URL https://aistudio.google.com/prompts/new_chat?model=gemini-2.0-flash-exp.

Yuriy Gorodnichenko, Tho Pham, and Oleksandr Talavera. The voice of monetary policy. *American Economic Review*, 113(2):548–84, February 2023a. doi: 10.1257/aer.20220129. URL <https://www.aeaweb.org/articles?id=10.1257/aer.20220129>.

Yuriy Gorodnichenko, Tho Pham, and Oleksandr Talavera. The voice of monetary policy. *American Economic Review*, 113(2):548–584, 2023b.

Stephen Hansen, Michael McMahon, and Matthew Tong. The long-run information effect of central bank communication. *Journal of Monetary Economics*, 108:185–202, 2019.

Adam Hayes, Michael J Boyle, and Jared Ecker. Purchasing power: What it means, how it’s measured, and examples. <https://www.investopedia.com/terms/p/purchasingpower.asp>, 2024. Accessed 7-May-2025.

Heartex. Label studio documentation, 2024. URL <https://labelstud.io/guide/>. Accessed: 2025-04-11.

Jens Hilscher, Kyle Nabors, and Alon Raviv. Information in central bank sentiment: An analysis of fed and ecb communication. *Available at SSRN 4797935*, 2024.

John Hooley, Mr Ashraf Khan, Claney Lattie, Istvan Mak, Ms Natalia Salazar, Amanda Sayegh, and Mr Peter Stella. *Quasi-fiscal implications of central bank crisis interventions*. International Monetary Fund, 2023.

Allen H. Huang, Hui Wang, and Yi Yang. Finbert: A large language model for extracting information from financial text. *Contemporary Accounting Research*, 40(2):806–841, 2023. doi: <https://doi.org/10.1111/1911-3846.12832>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/1911-3846.12832>.

Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, et al. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*, 2024.

Joaquín Iglesias, A Ortiz, and Tomasa Rodrigo. How do the emerging markets central bank talk? a big data approach to the central bank of turkey. Technical report, Working Paper, 2017.

411 Mehmet Ivrendi and Zekeriya Yildirim. Monetary Policy Shocks and Macroeconomic Variables: Evidence from
412 Fast Growing Emerging Economies [Dataset], 2013. URL <https://doi.org/10.7910/DVN/23957>.

413 Aparna Jayashankar and Anthony Murphy. High inflation disproportionately hurts low-income households. *Dal-*
414 *las Fed Economics*, January 2023. URL [https://www.dallasfed.org/research/economics/2023/](https://www.dallasfed.org/research/economics/2023/0110)
415 [0110](https://www.dallasfed.org/research/economics/2023/0110).

416 Narasimhan Jegadeesh and Di Wu. Deciphering fedspeak: The information content of fomc meetings. *Monetary*
417 *Economics: Central Banks–Policies & Impacts eJournal*, 2017.

418 Sophia Kazinnik and David H. Papell. Monetary policy rules in practice: The case of israel. *International Review*
419 *of Economics & Finance*, 76:308–320, 2021. ISSN 1059-0560. doi: <https://doi.org/10.1016/j.iref.2021.06.012>.
420 URL <https://www.sciencedirect.com/science/article/pii/S1059056021001386>.

421 Divya Kirti, Yan Liu, Maria Soledad Martinez Peria, Prachi Mishra, and Jan Strasky. Tracking economic
422 and financial policies during covid-19: An announcement-level database. IMF Working Paper 2022/114,
423 International Monetary Fund, Washington, DC, 2022.

424 Rohit Kumar, Sourabh Bikas Paul, and Nikita Singh. Words that move markets- quantifying the impact of rbi’s
425 monetary policy communications on indian financial market, 2024. URL [https://arxiv.org/abs/2411.](https://arxiv.org/abs/2411.04808)
426 [04808](https://arxiv.org/abs/2411.04808).

427 Aixin Liu, Bei Feng, Bing Xue, Bingxuan Wang, Bochao Wu, Chengda Lu, Chenggang Zhao, Chengqi Deng,
428 Chenyu Zhang, Chong Ruan, et al. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*, 2024.

429 Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke
430 Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach, 2019.

431 Shayne Longpre, Nikhil Singh, Manuel Cherep, Kushagra Tiwary, Joanna Materzynska, William Brannon,
432 Robert Mahari, Naana Obeng-Marnu, Manan Dey, Mohammed Hamdy, et al. Bridging the data provenance
433 gap across text, speech and video. *arXiv preprint arXiv:2412.17847*, 2024.

434 Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization, 2019. URL [https://arxiv.org/](https://arxiv.org/abs/1711.05101)
435 [abs/1711.05101](https://arxiv.org/abs/1711.05101).

436 Tim Loughran and Bill McDonald. When is a liability not a liability? textual analysis, dictionaries, and 10-ks.
437 *Journal of Finance*, 2010. URL <https://ssrn.com/abstract=1331573>. Forthcoming.

438 Christina Majaski, Khadija Khartit, and Jiwon Ma. Seasonal adjustment: Definition, calculation methods,
439 example. <https://www.investopedia.com/terms/s/seasonal-adjustment.asp>, 2022. Accessed
440 7-May-2025.

441 Rohin Manvi, Samar Khanna, Marshall Burke, David Lobell, and Stefano Ermon. Large language models are
442 geographically biased, 2024. URL <https://arxiv.org/abs/2402.02680>.

443 Puneet Mathur, Atula Neerkaje, Malika Chhibber, Ramit Sawhney, Fuming Guo, Franck Dernoncourt, Sang-
444 hamitra Dutta, and Dinesh Manocha. Monopoly: Financial prediction from monetary policy conference
445 videos using multimodal cues. In *Proceedings of the 30th ACM International Conference on Multimedia*,
446 pages 2276–2285, 2022.

447 Akira Matsui, Xiang Ren, and Emilio Ferrara. Using word embedding to reveal monetary policy explanation
448 changes. In Udo Hahn, Veronique Hoste, and Amanda Stent, editors, *Proceedings of the Third Workshop on*
449 *Economics and Natural Language Processing*, pages 56–61, Punta Cana, Dominican Republic, November
450 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.econlp-1.8. URL [https://](https://aclanthology.org/2021.econlp-1.8/)
451 aclanthology.org/2021.econlp-1.8/.

452 Filippo Occhino. Quantitative easing and direct lending in response to the covid-19 crisis. *Journal of Money,*
453 *Credit and Banking*, 2020.

454 OpenAI. Gpt-4.1, 2025. URL <https://openai.com/index/gpt-4-1/>.

455 Yusuke Oshima, Yoichi Matsubayashi, et al. Monetary policy communication of the bank of japan: Computa-
456 tional text analysis. *Graduate School of Economics Working paper*, 2018.

457 Huzaifa Pardawala, Siddhant Sukhani, Agam Shah, Veer Kejriwal, Abhishek Pillai, Rohan Bhasin, Andrew
458 DiBiasio, Tarun Mandapati, Dhruv Adha, and Sudheer Chava. Subjective-qa: Measuring subjectivity in
459 earnings call transcripts’ qa through six-dimensional feature analysis, 2025. URL [https://arxiv.org/](https://arxiv.org/abs/2410.20651)
460 [abs/2410.20651](https://arxiv.org/abs/2410.20651).

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library, 2019.

Marek Rozkrut, Krzysztof Rybiński, Lucyna Sztaba, and Radosław Szwaja. Quest for central bank communication: Does it pay to be “talkative”? *European Journal of Political Economy*, 23(1):176–206, 2007. ISSN 0176-2680. doi: <https://doi.org/10.1016/j.ejpoleco.2006.09.011>. URL <https://www.sciencedirect.com/science/article/pii/S0176268006000917>. Central bank transparency and central bank communication.

Troy Segal, Charles Potters, and Suzanne Kvilhaug. What is a central bank, and does the u.s. have one? <https://www.investopedia.com/terms/c/centralbank.asp>, 2023. Accessed 7-May-2025.

Agam Shah, Suvan Paturi, and Sudheer Chava. Trillion dollar words: A new financial dataset, task & market analysis. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6664–6679, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.acl-long.368. URL <https://aclanthology.org/2023.acl-long.368>.

Agam Shah, Abhinav Gullapalli, Ruchit Vithani, Michael Galarnyk, and Sudheer Chava. Finer-ord: Financial named entity recognition open research dataset, 2024. URL <https://arxiv.org/abs/2302.11157>.

Raj Shah, Kunal Chawla, Dheeraj Eidnani, Agam Shah, Wendi Du, Sudheer Chava, Natraj Raman, Charese Smiley, Jiaao Chen, and Diyi Yang. When FLUE meets FLANG: Benchmarks and large pretrained language model for financial domain. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 2322–2335, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.emnlp-main.148. URL <https://aclanthology.org/2022.emnlp-main.148>.

Ranking Sovereign Wealth Fund Institute. Central bank rankings, 2025. URL <https://www.swfinstitute.org/fund-rankings/central-bank>.

Raul Cruz Tadle. Fomc minutes sentiments and their impact on financial markets. *Journal of Economics and Business*, 118:106021, 2022. ISSN 0148-6195. doi: <https://doi.org/10.1016/j.jeconbus.2021.106021>. URL <https://www.sciencedirect.com/science/article/pii/S0148619521000394>.

Nikita Tatarinov, Siddhant Sukhani, Agam Shah, and Sudheer Chava. Language modeling for the future of finance: A quantitative survey into metrics, tasks, and data opportunities, 2025. URL <https://arxiv.org/abs/2504.07274>.

Ellen Tobback, Stefano Nardelli, and David Martens. *Between hawks and doves: measuring central bank communication*. ECB Working paper, 2017.

USCongress. Nomination hearing for jared bernstein to be chairman of the council of economic advisers. <https://www.govinfo.gov/content/pkg/CHRG-118shrg57489/html/CHRG-118shrg57489.htm>, 2023. U.S. Government Publishing Office, 118th Congress, 1st Session.

Marco Vega and Erick Lahura. Assessing central bank communication through monetary policy statements: Results for colombia, chile and peru. Technical report, Banco Central de Reserva del Perú, 2020.

Benjamin Warner, Antoine Chaffin, Benjamin Clavié, Orion Weller, Oskar Hallström, Said Taghadouini, Alexis Gallagher, Raja Biswas, Faisal Ladhak, Tom Aarsen, Nathan Cooper, Griffin Adams, Jeremy Howard, and Iacopo Poli. Smarter, better, faster, longer: A modern bidirectional encoder for fast, memory efficient, and long context finetuning and inference, 2024. URL <https://arxiv.org/abs/2412.13663>.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander Rush. Transformers: State-of-the-art natural language processing. In Qun Liu and David Schlangen, editors, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online, October 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.emnlp-demos.6. URL <https://aclanthology.org/2020.emnlp-demos.6>.

Alfred Y-T. Wong and Tom Pak Wing Fong. Analysing interconnectivity among economies. *Emerging Markets Review*, 12(4):432–442, 2011. ISSN 1566-0141. doi: <https://doi.org/10.1016/j.ememar.2011.06.004>. URL <https://www.sciencedirect.com/science/article/pii/S156601411100046X>.

- 514 Qianqian Xie, Weiguang Han, Xiao Zhang, Yanzhao Lai, Min Peng, Alejandro Lopez-Lira, and Jimin Huang.
515 Pixiu: A large language model, instruction data and evaluation benchmark for finance, 2023. URL <https://arxiv.org/abs/2306.05443>.
516
- 517 Yang Zhang, Wenbo Yang, Jun Wang, Qiang Ma, and Jie Xiong. Camef: Causal-augmented multi-modality event-
518 driven financial forecasting by integrating time series patterns and salient macroeconomic announcements,
519 2025. URL <https://arxiv.org/abs/2502.04592>.
- 520 HongFang Zhou, Jin Li, JunHuai Li, FaCun Zhang, and YingAn Cui. A graph clustering method for community
521 detection in complex networks. *Physica A: Statistical Mechanics and Its Applications*, 469:551–562, 2017.

522 Contents

523	A Release and Responsibility Statements	18
524	A.1 Author Statement	18
525	A.2 Hosting, Licensing and Maintenance	18
526	A.3 Disclaimer	19
527	A.4 Ethical Considerations	20
528	A.5 Dataset Clause and Terms of Use	20
529	A.6 Scraping Policy Consideration	21
530	B Extended Data Construction Process	22
531	B.1 Acronyms for Bank Names	22
532	B.2 Bank Selection	22
533	B.3 Dataset Construction and Cleaning	24
534	B.4 Dataset Statistics	25
535	B.5 Metadata Reporting	26
536	B.6 Instructions Given	27
537	B.6.1 Initial Research	27
538	B.6.2 Sample Annotation Task	27
539	B.6.3 Annotation Guide Task	28
540	B.6.4 Annotation Task	28
541	B.6.5 Annotation Disagreement Resolution	29
542	B.7 Annotation Interface	29
543	B.8 Why Unique Annotation Guides?	30
544	B.8.1 Manual Cross Validation	30
545	B.8.2 4-Gram Semantic Overlap	30
546	B.9 Why Irrelevant Label for Stance?	30
547	C Models and Benchmarking	33
548	C.1 Prompts for Benchmarking	33
549	C.1.1 Zero-Shot Prompt: Stance Detection	33
550	C.1.2 Zero-Shot Prompt: Temporal Classification	34
551	C.1.3 Zero-Shot Prompt: Uncertainty Estimation	35
552	C.1.4 Prompting Structure for Few-Shot prompting	36
553	C.1.5 Prompting Structure for prompting with Annotation Guides	37
554	C.2 Model Details	37
555	C.3 Model Specifications, Compute, and Providers	38
556	C.4 Results	39
557	C.4.1 General (All-Banks) Setup Models	39
558	C.4.2 Specific Bank Setup Models	42
559	C.5 Few shot & Annotation Guide Few Shots Prompting	45

560	C.6 Whole Model Performance Gain	45
561	C.6.1 Qualitative Reasoning	45
562	C.6.2 Statistical Reasoning	46
563	C.7 Experiment: Training Set Size	48
564	C.8 Hyperparameters	48
565	C.8.1 WCB Models	48
566	C.8.2 Specific Bank Models	48
567	D Extended Experiments and Analysis	51
568	D.1 People’s Bank of China LLM Error Analysis	51
569	D.1.1 Stance Detection	51
570	D.1.2 Temporal Classification	51
571	D.1.3 Uncertainty Estimation	51
572	D.2 Generating Meeting Minutes	51
573	D.2.1 Methodology	51
574	D.2.2 Hawkishness Measure	52
575	D.3 Human Evaluation	54
576	D.4 Transfer Learning	54
577	D.4.1 Beyond 25 Banks: Czech National Bank	54
578	D.4.2 Beyond Finance: The Future of Conservation Climate, Forestry, and Natural	
579	Resources	55
580	E Extended Economic Analysis	58
581	E.1 Data Collection and Aggregation	58
582	E.2 Hawkishness and Inflation	59
583	E.3 Hindsight 20/20, Foresight Uncertain	72
584	F Extended Related Works	73
585	G Glossary	74
586	H Central Banks Information, Annotator Disagreement, and Annotation Guides	75
587	H.1 Federal Open Market Committee	75
588	H.2 People’s Bank of China	81
589	H.3 Bank of Japan	86
590	H.4 Bank of England	93
591	H.5 Swiss National Bank	99
592	H.6 Central Bank of Brazil	104
593	H.7 Reserve Bank of India	110
594	H.8 European Central Bank	116
595	H.9 Central Bank of the Russian Federation	122
596	H.10 Central Bank of China (Taiwan)	129

597	H.11 Monetary Authority of Singapore	135
598	H.12 Bank of Korea	140
599	H.13 Reserve Bank of Australia	147
600	H.14 Bank of Israel	154
601	H.15 Bank of Canada	159
602	H.16 Bank of Mexico	164
603	H.17 Narodowy Bank Polski	169
604	H.18 Central Bank of Turkey	175
605	H.19 Bank of Thailand	180
606	H.20 Central Bank of Egypt	186
607	H.21 Bank Negara Malaysia	192
608	H.22 Central Bank of the Philippines	198
609	H.23 Central Bank of Chile	203
610	H.24 Central Reserve Bank of Peru	209
611	H.25 Bank of the Republic	216
612	H.26 Czech National Bank	222

A Release and Responsibility Statements

A.1 Author Statement

We confirm that the WCB dataset is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC-BY-NC-SA 4.0) license, which allows others to share, copy, distribute, and transmit the work, as well as to adapt the work, provided that appropriate credit is given for a non-commercial use, a link to the license is provided, and any changes made are indicated. Additionally, the derived work needs to be made available under a similar license.

A.2 Hosting, Licensing and Maintenance

The **World Central Banks (WCB)** dataset is distributed under the *Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International* (CC-BY-NC-SA 4.0) license. This license allows users to share, adapt, and build upon the dataset for non-commercial purposes, provided proper attribution is given and derivative works are shared alike.

To ensure broad accessibility and reproducibility, we host the dataset, codebase, models, and visualizations across three major platforms:

The WCB dataset and resources are available at:

GitHub:

<https://github.com/gtfintechlab/WorldCentralBanks>

Hugging Face Collection (Models + Datasets):

https:

[//huggingface.co/collections/gtfintechlab/wcb-678965e38178c63158b45fdf](https://huggingface.co/collections/gtfintechlab/wcb-678965e38178c63158b45fdf)

Website:

<https://gcb-web-bb21b.web.app/>

The Hugging Face collection includes:

- **Four datasets:**

- A unified corpus of 380,200 scraped and cleaned sentences
- An annotated dataset of 25,000 sentences across three tasks
- A Czech National Bank subset (CNB)
- Bank-specific subsets for 25 central banks

- **78 trained models:**

- 75 bank-specific models (trained on 3 tasks for each of the 25 central banks)
- 3 models trained on the general dataset (trained on 3 tasks using the aggregated dataset of all central banks)

The links to the 1k annotated sentences for all 25 central banks, their corresponding Bank-Specific Setup models, the aggregated 25k annotated sentences dataset, and the three General Setup models are provided in Table 3.

The GitHub repository offers comprehensive access to:


























- All raw, cleaned, and sanitized data
- Structured metadata
- Complete benchmarking and training pipelines
- Instructions to reproduce results from benchmarking as well as additional experiments

Each folder contains the policy data for each of the 25 banks. Each bank’s folder contains the corresponding data for each year. Furthermore, in the sanitized data folder, each of the banks contains

a **Metadata** JSON file. The metadata JSON as seen in Figure 3 Appendix B.5 is stored as a large dictionary where the doc_id is mapped to the unique documents within a specific bank for a specific year. In addition, a global **Master Metadata JSON** is present, which is an aggregate dictionary mapping for the entire dataset - all the bank specific metadata JSON files compiled into a large *master* file with the same format. More information on the metadata is reported in B.5.

The WCB website complements these resources with dynamic dashboards showcasing model performance and interactive visualizations.

Table 3: The datasets for each central banks and best performing models for each task (Stance Detection, Temporal Classification, and Uncertainty Estimation) are available hyper-linked in the table below. The specific hyperparameter used for each model are summarized in Tables 20, 21, 22 in Appendix C.8.

Bank	Dataset	Model		
		Stance	Temporal	Uncertainty
General Setup	Dataset	Model	Model	Model
 Federal Open Market Committee	Dataset	Model	Model	Model
 People’s Bank of China	Dataset	Model	Model	Model
 Bank of Japan	Dataset	Model	Model	Model
 Bank of England	Dataset	Model	Model	Model
 Swiss National Bank	Dataset	Model	Model	Model
 Central Bank of Brazil	Dataset	Model	Model	Model
 Reserve Bank of India	Dataset	Model	Model	Model
 European Central Bank	Dataset	Model	Model	Model
 Central Bank of the Russian Federation	Dataset	Model	Model	Model
 Central Bank of China	Dataset	Model	Model	Model
 Monetary Authority of Singapore	Dataset	Model	Model	Model
 Bank of Korea	Dataset	Model	Model	Model
 Reserve Bank of Australia	Dataset	Model	Model	Model
 Bank of Israel	Dataset	Model	Model	Model
 Bank of Canada	Dataset	Model	Model	Model
 Bank of Mexico	Dataset	Model	Model	Model
 National Bank of Poland	Dataset	Model	Model	Model
 Central Bank of Turkey	Dataset	Model	Model	Model
 Bank of Thailand	Dataset	Model	Model	Model
 Central Bank of Egypt	Dataset	Model	Model	Model
 Bank Negara Malaysia	Dataset	Model	Model	Model
 Central Bank of the Philippines	Dataset	Model	Model	Model
 Central Bank of Chile	Dataset	Model	Model	Model
 Central Reserve Bank of Peru	Dataset	Model	Model	Model
 Bank of the Republic	Dataset	Model	Model	Model

A.3 Disclaimer

We provide the following two disclaimers regarding our work.

Geopolitical Disclaimer All references to country names, territories, flags, boundaries, or other geopolitical identifiers in this work are drawn solely from publicly available sources, principally datasets and publications produced by international organizations such as the International Monetary Fund (IMF), the World Bank, and the United Nations. Their inclusion is solely for descriptive or analytical purposes.

Neither the authors nor their affiliated institution(s):

- Endorse or oppose any political stance, territorial claim, or status implied by these identifiers;
- Express any opinion concerning the legal status of any country, territory, city, or area, or the delimitation of its frontiers or boundaries; or

- Imply recognition, approval, or disapproval of the policies, positions, or governments of the entities mentioned.

Financial Disclaimer The analyses and interpretations presented in this paper are provided only for academic discussion. They do not constitute official economic policy advice, legal guidance, or a recommendation to implement specific actions. Readers should exercise their own judgment before making decisions based on the information herein.

A.4 Ethical Considerations

In conducting our research, we adhere to strict ethical and legal policies and do not identify any risks prevalent in the research carried out. While our training process was designed to minimize the introduction of any bias, we acknowledge that residual bias may persist due to pre-existing biases in our data sources. However, we recognize that there may still be residual bias likely caused by pre-existing biases in the data sources. Furthermore, we acknowledge that our research is subject to certain limitations and potential biases inherent in our annotation and study process.

- **Educational Bias** Most of the annotators came from STEM fields, and this could have impacted the annotation process.
- **Geographic Bias** NLP datasets often exhibit geographical biases, due to disproportionately sourced data from specific regions, cultural contexts, and languages [Tatarinov et al. \[2025\]](#). In our research, we sourced our data from each individual central bank’s official website to minimize disproportion as much as possible.
- **Data Ethics** All the names, data, and views are sourced exclusively from the official website of the central banks and the World Bank. All data were obtained ethically and legally. Some central banks restrict access based on geographical factors, so to align with ethical research practices, we did not collect data from those central banks. As a result, our study is limited to 25 central banks whose data can be ethically accessed. We also included an additional bank, the Czech National Bank, whose data was accessed ethically and utilized exclusively for ablation studies.
- **Annotation Ethics** The annotation of the dataset was completed ethically. All annotation work was completed on a voluntarily basis, and annotators were made aware of the tasks and purpose of the project in advance. None of the annotators were paid to do the annotations.
- **Publicly Available Data** All the data utilized for this research will be made publicly available, along with the appropriate licenses under which it may be shared.
- **Language Model Ethics** All the language models employed for this research are publicly available and we complied with their license categories for our intended purposes. To curb the carbon footprint associated with large-scale pre-training, we only fine-tuned existing PLMs and ran prompt-based evaluations on LLMs. We cited all the models used.
- **Hyperparameter Reporting** All hyperparameter used in training the models are specified within Appendix C.8, thus making our model setup transparent and providing detailed information on our training process.
- **Model Ethics** Responsible AI practices guided our use of both PLMs and LLMs, including transparency, fairness, and reproducibility.

The research team is committed to promoting transparency, fairness, and accessibility by openly acknowledging any limitations in the findings, promoting ethical integrity, and encouraging responsible research practices.

A.5 Dataset Clause and Terms of Use

Purpose The Dataset is provided for the purpose of research and educational use in the field of NLP, Machine Learning, and related areas; and can be used to further the development, evaluation, or benchmarking of AI.

Usage Restriction Users of the Dataset should adhere to the terms of use for specific models or tasks. This includes respecting any limitations or use case prohibitions set forth by the original model’s and guide’s creators.

717 **Content Warning** The Dataset comprises of raw monetary policy filings from 25 central banks
718 spanning from 1996 to 2024. These documents may reflect sensitive economic positions, institutional
719 biases, or historically outdated policy perspectives. Therefore, users should exercise caution by
720 applying appropriate filtering, contextual analysis, and moderation protocols when using this dataset
721 for model training or decision-making tasks to ensure adherence to ethical standards.

722 **No Endorsement of Content** The conversations and data within this Dataset do not reflect the
723 views or opinions of the Dataset creators or any affiliated institutions. The dataset is provided as a
724 neutral resource for financial and economic research and should not be construed as endorsing any
725 specific viewpoints.

726 **Limitations of Liability** The authors and founders of this Dataset will not be liable for any claims,
727 damages, or other liabilities arising from the use of the dataset, including but not limited to the misuse,
728 interpretation, or reliance on any data contained within.

729 **A.6 Scraping Policy Consideration**

730 The data gathered for this research were collected using legal and ethical practices. We have ensured
731 that all of our datasets were gathered in full compliance with each institution's ethical requirements.
732 Specifically, for each bank we reviewed their official Terms of Use or Disclaimer page on their
733 websites and also inspected their site's `robots.txt` (if available) file to confirm that automated
734 crawling of public press releases and meeting-minutes pages was permitted. Care was taken to respect
735 access rules, avoid disallowed areas, and maintain server load limits through crawl delays. Therefore,
736 our dataset respects each bank's data-use policy.

737 B Extended Data Construction Process

738 B.1 Acronyms for Bank Names

739 We utilize the following acronyms within this paper in order to present our results in a clear and concise manner. The acronyms are based on the official abbreviation or standard usage.

Table 4: List of central banks and their corresponding acronyms.

Bank Name	Acronym
 Federal Open Market Committee	FOMC
 People’s Bank of China	PBoC
 Bank of Japan	BoJ
 Bank of England	BoE
 Swiss National Bank	SNB
 Central Bank of Brazil	BCB
 Reserve Bank of India	RBI
 European Central Bank	ECB
 Central Bank of the Russian Federation	CBR
 Central Bank of China (Taiwan)	CBCT
 Monetary Authority of Singapore	MAS
 Bank of Korea	BoK
 Reserve Bank of Australia	RBA
 Bank of Israel	BoI
 Bank of Canada	BoC
 Bank of Mexico	BdeM
 Narodowy Bank Polski	NBP
 Central Bank of Turkey	CBRT
 Bank of Thailand	BoT
 Central Bank of Egypt	CBE
 Bank Negara Malaysia	BNM
 Central Bank of the Philippines	BSP
 Central Bank of Chile	CBoC
 Central Reserve Bank of Peru	BCRP
 Bank of the Republic	BanRep

740

741 B.2 Bank Selection

742 We had a particular set of criteria for the selection of the central banks within our corpus. We describe
 743 the reason for not selecting certain banks in Table 5. Banks/Economies that fall under the jurisdiction
 744 of the European Central Bank were excluded from this dataset as their monetary policy is governed
 745 by the ECB and they use a standardized currency (the Euro). We observed that certain documents
 746 from the Bank of Chile in 2022 were in Spanish with no official translation and 2018 documents from
 747 the Central Bank of the Republic of Turkey were unavailable.

748 **Community Involvement** After extracting and cleaning the data, we select 28 central banks,
 749 including the 26 within our analysis and two additional banks: Sveriges Riksbank and Bank Indonesia.
 750 The Czech National Bank was used as a hold-out dataset, with only 500 sentences evaluated. Out of
 751 the corpus of cleaned data for 28 banks, we randomly sample the latter two banks (Sveriges Riksbank
 752 and Bank Indonesia) that we open to the community of AI, NLP, and economic researchers to analyze
 753 and annotate. This step aims to increase engagement of our work within the community.

Table 5: List of excluded central banks and reasons for their exclusion. Rows marked with * fall under the jurisdiction of the European Central Bank, as stated by [European Central Bank \[2025\]](#).

Bank	Reason for Exclusion
Bank of Finland	Bank falls under jurisdiction of European Central Bank*
Bank of France	Bank falls under jurisdiction of European Central Bank*
Bank of Greece	Bank falls under jurisdiction of European Central Bank*
Bank of Italy	Bank falls under jurisdiction of European Central Bank*
Bank of Portugal	Bank falls under jurisdiction of European Central Bank*
Bank of Spain	Bank falls under jurisdiction of European Central Bank*
Central Bank of Ireland	Bank falls under jurisdiction of European Central Bank*
Central Bank of Luxembourg	Bank falls under jurisdiction of European Central Bank*
Deutsche Bundesbank	Bank falls under jurisdiction of European Central Bank*
National Bank of Belgium	Bank falls under jurisdiction of European Central Bank*
National Bank of the Republic of Austria	Bank falls under jurisdiction of European Central Bank*
Netherlands Bank	Bank falls under jurisdiction of European Central Bank*
Saudi Central Bank	Documents inaccessible at scale
Central Bank of Iran	Documents inaccessible at scale
Central Bank of Iraq	Documents inaccessible at scale
Norges Bank	Lack of consistent meeting minute level equivalent document
Central Bank of Lebanon	Lack of consistent meeting minute level equivalent document
National Bank of Denmark	Lack of consistent meeting minute level equivalent document
State Bank of Vietnam	Lack of consistent meeting minute level equivalent document
Qatar Central Bank	Lack of consistent meeting minute level equivalent document
Bank of Algeria	No meeting minute level document
Hong Kong Monetary Authority	No meeting minute level document
Central Bank of Nigeria	No meeting minute level document
Central Bank of UAE	No meeting minute level document
Central Bank of Argentina	Insufficient number of documents

B.3 Dataset Construction and Cleaning

Extraction Based on our selection of the central banks as highlighted in Table 6, we scraped the central banks’ websites for the necessary documents. We targeted meeting minute documents or their closet equivalent in terms of content and formatting. The documents for each central bank alongside the link we used to find these documents can be found in Table 6. For the scraping, we used the Python library `selenium` to navigate through the pages and download the data. It is important to note that some central banks’ data was in PDF format whilst some were just text files. This led to a split in the data wherein certain banks’ data was downloaded as a PDF and converted to Markdown by using the `PyMuPDF4LLM` Python library. If the data was in text format, it was kept as is and any Microsoft Word documents (Docx) were converted to text.

Cleaning and Normalizing Once all the documents were downloaded and sorted based on year, we began cleaning the documents by removing everything apart from the actual sentences of the document. This involved removing any charts and other visuals, indiscernible characters, phone numbers, email addresses, and other such unnecessary information that may have been scraped.

After the data was cleaned, the sentences for each central bank were consolidated into separate JSON files, with the years as keys and the list of the sentences as the values. Each sentence was then normalized by stripping the whitespace and removing any special characters. A sampling script was then used to randomly sample 1000 sentences from each central bank, with equal weight being given to each year. The final dataset consisted of 1000 normalized sentences from each central bank. This process enabled us to have a clean and rich dataset that was ready to annotate.

Table 6: The central banks that were chosen alongside their type of document and format. Each central bank name is hyperlinked to the source of that central bank’s data.



















































Bank	Type of Document	Format
 Federal Reserve System	Monetary Policy Minutes	Text
 People’s Bank of China	Monetary Policy Minutes	Text
 Bank of Japan	Monetary Policy Minutes	PDF & Text
 Bank of England	Monetary Policy Minutes	PDF
 Swiss National Bank	Monetary policy assessments	PDF
 Central Bank of Brazil	Monetary Policy Minutes	PDF
 Reserve Bank of India	Monetary Policy Minutes	PDF
 European Central Bank	Monetary policy accounts	Text
 Central Bank of the Russian Federation	Key rate decisions	Text
 Central Bank of China (Taiwan)	Monetary Policy Minutes	PDF & Docx
 Monetary Authority of Singapore	Monetary policy statements	Text
 Bank of Korea	Monetary Policy Minutes	PDF
 Reserve Bank of Australia	Monetary Policy Minutes	Text
 Bank of Israel	Monetary Policy Minutes	PDF
 Bank of Canada	Monetary Policy Press releases	Text
 Bank of Mexico	Monetary Policy Minutes	PDF
 National Bank of Poland	Monetary Policy Minutes	PDF
 Central Bank of Turkey	Monetary Policy Meeting summary	Text
 Bank of Thailand	Monetary Policy Minutes	PDF
 Central Bank of Egypt	Monetary Policy Press releases	PDF
 Bank Negara Malaysia	Monetary policy statements	Text
 Central Bank of the Philippines	Monetary Board highlights	PDF
 Central Bank of Chile	Monetary Policy Minutes	PDF
 Central Reserve Bank of Peru	Monetary Policy Statements	PDF
 Bank of the Republic (Colombia)	Monetary Policy Minutes	PDF & Text

Table 7: Statistics for various central banks over different year ranges. The ‘*’ indicates a corpus-wide average taken over the whole dataset.

Bank	Year Range	Years	Total			Average		
			Sentences	Characters	Words	Chars/Sent	Words/Sent	Sents/Year
 FOMC	1996–2024	29	47,740	8,784,484	1,331,461	184.01	27.89	1,646.21
 PBoC	2002–2024	21	1,189	264,511	39,082	222.47	32.87	56.62
 BoJ	1998–2024	27	70,941	13,854,901	2,174,537	195.30	30.65	2,627.44
 BoE	1997–2024	28	68,281	10,136,875	1,642,593	148.46	24.06	2,438.61
 SNB	2001–2024	24	3,377	369,886	59,585	109.53	17.64	140.71
 BCB	2000–2024	24	11,623	2,086,611	316,659	179.52	27.24	484.29
 RBI	2016–2024	9	11,102	1,613,434	258,117	145.33	23.25	1,233.56
 ECB	2015–2024	10	21,044	3,734,051	582,948	177.44	27.70	2,104.40
 CBR	2014–2024	10	3,430	433,937	68,111	126.51	19.86	343.00
 CBC	2017–2024	8	5,317	955,055	148,257	179.62	27.88	664.62
 MAS	2001–2024	24	2,183	300,453	47,987	137.63	21.98	90.96
 BoK	2010–2024	15	11,120	2,592,561	403,395	233.14	36.28	741.33
 RBA	2006–2024	19	25,544	3,912,327	619,701	153.16	24.26	1,344.42
 BoI	2006–2024	18	13,199	2,280,543	372,816	172.78	28.25	733.28
 BoC	2011–2024	14	1,499	197,542	31,575	131.78	21.06	107.07
 BdeM	2018–2024	7	14,290	2,450,950	377,409	171.52	26.41	2,041.43
 NBP	2007–2024	18	11,388	2,101,567	334,635	184.54	29.38	632.67
 CBRT	2015–2024	10	9,801	1,457,271	221,108	148.69	22.56	980.10
 BoT	2011–2024	14	7,828	1,290,443	191,807	164.85	24.50	559.14
 CBE	2005–2024	20	2,369	399,815	62,496	168.77	26.38	118.45
 BNM	2004–2024	21	2,344	311,502	46,766	132.89	19.95	111.62
 BSP	2002–2024	23	15,697	2,634,059	411,611	167.81	26.22	682.48
 CBoC	2018–2024	7	3,907	657,712	104,124	168.34	26.65	558.14
 BCRP	2001–2024	24	4,199	695,774	111,294	165.70	26.50	174.96
 BanRep	2008–2024	17	10,788	2,054,048	331,089	190.40	30.69	634.59
WCB	1996–2024	17.64*	380,200	65,570,312	10,289,163	172.46*	27.06*	13,110.34*

B.5 Metadata Reporting

We format our metadata for each central bank and the corpus of 380,200 sentences as a whole in a standardized format. Our metadata is reported in both CSV and JSON format on our GitHub and HuggingFace for each central bank’s dataset. The metadata reported is for all sentences for a given bank rather than just the annotated sentences. The formatting of the metadata in JSON is shown below in Figure 3. The metadata in CSV format is similarly formatted, with each row containing the following columns: sentence, bank, year, original_document_name, release_date, start_date, end_date, minutes_link, cleaned_document_name, doc_id, sentence_length.

```
Metadata JSON formatting

{
  "central_bank_name": {
    "year": {
      "document_key": {
        "release_date": "DD-MM-YYYY",
        "start_date": "DD-MM-YYYY",
        "end_date": "DD-MM-YYYY",
        "minutes_link": "URL to the source document",
        "cleaned_document_name": "Filename of cleaned document",
        "original_document_name": "Filename of original document",
        "sentences": [
          "First sentence from the document.",
          "Second sentence from the document.",
          "...
        ]
      }
    }
  }
}
```

Figure 3: Template of JSON-style formatting for the metadata of each central bank in the corpus.

Central Bank Document Limitations and Observations

- **People’s Bank of China:** Many documents display vague dates like “days ago” or “several days ago.” In such cases, we defaulted to using the release date as the meeting date.
- **Central Reserve Bank of Peru:** There is a typographical error in the document dated July 6, 2002. Another document intended for March 7, 2002 is incorrectly labeled as March 7, 2001 on the central bank’s website.
- **Bank of Russia:** Two documents’ names are written in Russian but the actual sentences are in english. Their filenames have been manually renamed to maintain consistency. The two files were [document 1](#) and [document 2](#).
- **Central Bank of Chile:** Some document filenames include the start date, but the actual meeting end date may be one day later. In rare cases, the date in the filename corresponds to the end date instead. These were handled accordingly. Additionally, a document from May 2022 is protected by Microsoft, which prevented access.
- **Bank of Israel:** No meeting documents were found for 2017, only a general report. In many cases, the link date and actual discussion date do not align clearly. For example:
“Report to the public of the Bank of Israel’s discussions prior to setting the interest rate for December 2006. The discussions took place on November 26, 2006.”
This makes automated date extraction unreliable due to lack of consistent formatting.
- **Central Bank of Egypt:** A document from March 28, 2004 was found to be irrelevant as a whole and thus excluded from processing. The document only mentioned the change of date of the meeting and nothing more.

805 • **Central Bank of Brazil:** One document (MIN200044-COPOM20000215-44th COPOM
806 Minutes) from February 15, 2000 has a release date earlier than the end date, which
807 is an anomaly. A document from the 49th meeting in 2000 is too noisy and unreadable using
808 standard PDF libraries and was therefore excluded.

809 B.6 Instructions Given

810 **Terms of Participation** The annotators in this project were volunteers or students working toward
811 research credits at the Georgia Institute of Technology. Individuals who consistently demonstrated
812 high-quality contributions were awarded co-authorship on this paper. Participants may distribute their
813 own work but must not reference the lab or affiliates without prior consent. All external distributions
814 must comply with applicable intellectual property rights and data sharing agreements.

815 The following subsections describe the instructions provided to participants based on the tasks
816 assigned.

817 B.6.1 Initial Research

818 The participants were tasked with writing a summary of their assigned central bank to help shape
819 the initial annotation guideline. Annotators were expected to read several minutes or equivalent
820 documents from Monetary Policy Meetings of their assigned central bank in order to gain insight into
821 the intended structure and content of the annotation guideline. The annotators were instructed to cite
822 all sources used and to not refer to any third-party data sources for their research.

823 Researching Central Bank

- 824 1. **Summary of the Monetary Policy:** Summarize the Monetary Policy of the central bank.
825 This should be about 2-3 paragraphs long briefly mentioning the central bank's money
826 supply, interest rates, and a description of the central bank's key historical events that have
827 shaped the policy, as well as any other indicators. Determine whether the central bank claims
828 a specific mandate. For example, the FOMC has a dual mandate: maximum employment
829 and price stability.
- 830 2. **Numerical:** List the key numbers mentioned in the monetary policy such as Inflation rate,
831 Interest rate, and so on. Be as specific as you can, providing dates with each number.
- 832 3. **Trends:** Try to identify any trends within the monetary policy. This will require you to
833 research your respective central bank's policy over the last 5 to 10 years.
- 834 4. **Impact of Events:** Use a few lines to explain the impact of various global events such as
835 the COVID pandemic and the 2008 crisis on your specific central bank's monetary policy.
836 Additionally, observe any impacts that domestic events have had on the monetary policy.
837 (For example, demonetization for India).

838 Guidelines Provided to Annotators

- 839 • It is of utmost importance that you **do not not communicate** with your peers regarding this
840 task related to your central bank.
- 841 • The annotation guide should be written in clear and concise language. It must be in a PDF
842 format.

843 B.6.2 Sample Annotation Task

844 Each annotator was expected to complete the 100 assigned sentences and become familiar with the
845 Label Studio interface (Figure 4 Appendix B.7). Annotators followed the instructions for annotations.
846 These instructions were also made available on Label Studio. The annotators were told that each
847 sentence must be annotated along the following three dimensions:

- 848 • **Stance Detection:** Hawkish, Dovish, Neutral, or Irrelevant.
- 849 • **Temporal Classification:** Forward Looking or Not Forward Looking.
- 850 • **Uncertainty Estimation:** Certain or Uncertain.

851 Examples for Each Dimension

852 1. **Annotations:** Annotate the given data according to the three dimensions. Each sentence
853 will have these three dimensions associated with it.

- 854 • **Stance Detection:** Hawkish, Dovish, Neutral, or Irrelevant
- 855 – "The Committee then turned to a discussion of the economic and financial outlook,
856 the ranges for the growth of money and debt in 1996, and the implementation of
857 monetary policy over the intermeeting period ahead." Neutral
- 858 – "To support the Committee's decision to raise the target range for the federal
859 funds rate, the Board of Governors voted unanimously to raise the interest rates on
860 required and excess reserve balances to 2.": Hawkish
- 861 – "Inflation has moderated since earlier in the year, and longer-term inflation expecta-
862 tions have remained stable.": Dovish
- 863 – "It was agreed that the next meeting of the Committee would be held on Tuesday,
864 August 21, 2001.": Irrelevant
- 865 • **Temporal Classification:** Forward Looking or Not Forward Looking
- 866 – "Although the pace of economic recovery is likely to be moderate for a time, the
867 Committee anticipates a gradual return to higher levels of resource utilization in a
868 context of price stability.": Forward Looking
- 869 – "Commodity prices had been mixed recently after trending down earlier.": Not
870 Forward Looking
- 871 • **Uncertainty Estimation:** Certain or Uncertain.
- 872 – "To support the Committee's decision to raise the target range for the federal funds
873 rate, the Board of Governors of the Federal Reserve System voted unanimously to
874 raise the interest rate paid on reserve balances to 0.": Certain
- 875 – "Some participants were concerned that inflation could rise as the recovery con-
876 tinued, and some business contacts had reported that producers expected to see an
877 increase in pricing power over time.": Uncertain

878 Guidelines Provided to Annotators

- 879 • It is of utmost importance that you **do not communicate** with your peers regarding the
880 annotations related to your assigned countries.
- 881 • The annotations must be your own work. Progress will be monitored through Label Studio.

882 B.6.3 Annotation Guide Task

883 The goal of this task was to create an annotation guide that teams would utilize for the final round of
884 annotations. Specifically, the annotators were tasked with generating a guide describing the various
885 labels that the sentences were to be annotated with. The annotators were instructed to specify and
886 justify what the labels mean for their central bank and were encouraged to consider their initial
887 annotation process and the first annotation guide while building this.

888 Guidelines Provided for Annotators

- 889 • The submission must be your own work, but you are welcome to work in teams and use
890 LLMs.
- 891 • Provide both the initial and final annotation guide.

892 B.6.4 Annotation Task

893 Through the Label Studio interface, the annotators were tasked with annotating each sentence assigned
894 to them on the following three tasks:

- 895 1. Stance Detection: Hawkish, Dovish, Neutral, Irrelevant
- 896 2. Temporal Classification: Forward Looking, Not Forward Looking
- 897 3. Uncertainty Estimation: Certain, Uncertain

898 B.6.5 Annotation Disagreement Resolution

899 The goal of this task was to resolve disagreements in the final annotations using Label Studio. The
900 annotators were teamed up in pairs.

901 The following instructional steps were provided to the annotators for resolving disagreements in the
902 annotations:

- 903 1. Collaborate with the other annotator on your project. Strictly adhere to the latest annotation
904 guide for your assigned central bank and refer to it while resolving disagreements.
- 905 2. Open the data manager in your project and review annotations where there was a dis-
906 agreement (where annotator agreement falls below 100%). These can be identified in the
907 “Agreement” column in the data manager.
- 908 3. Ensure that the entire sentence is selected for all annotations. If there is a disagreement,
909 the annotator with the correct annotation should keep it. For example, if annotator A and
910 annotator B disagree and annotator A’s annotation is correct, then annotator B should accept
911 A’s annotation, and annotator A should fix and accept B’s annotations as corrected.
- 912 4. Confirm that all annotations have been accepted and that each annotator’s name is marked
913 with a tick in the “Annotators” column in the data manager, indicating completion.

914 B.7 Annotation Interface

915 Tile one in Figure 4 shows the Label Studio interface presented to annotators during the initial
916 labeling phase. Annotators were required to select one label per category without knowledge of their
917 partner’s annotations. Tile 2 of Figure 4 shows the interface used during the review and disagreement
918 resolution phase, where annotators could view disagreement statistics and their partners labels.
919 During this phase, they were allowed to discuss, edit their partner’s labels, or accept them directly,
920 which are displayed in tiles 3a, 3b, and 3c of Figure 4. All annotations and review were completed on
921 the same platform Heartex [2024] to ensure consistency.

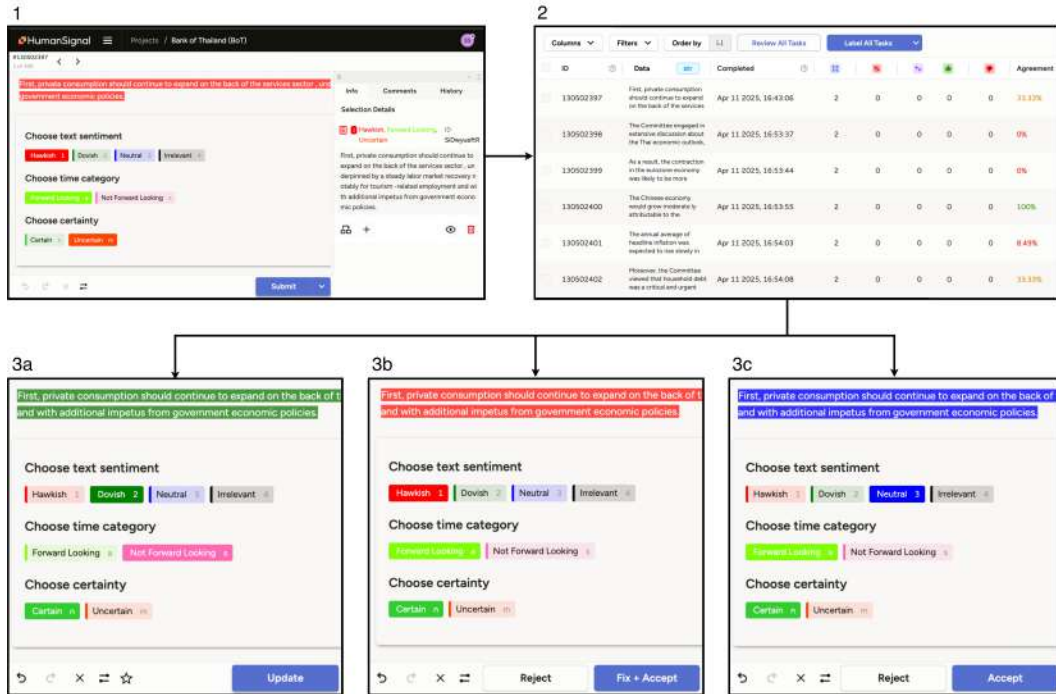


Figure 4: Tile 1 displays the interface to annotate sentences. Tile 2 displays the “Reviewer” view, allowing annotators to view disagreements after their initial annotations are complete. Tiles 3a, 3b, and 3c display the process of updating the annotation after pairs have reached a consensus.

B.8 Why Unique Annotation Guides?

B.8.1 Manual Cross Validation

To assess whether each central bank requires its own annotation guide, we conducted an experiment in which we first drew a random sample of 500 sentences from the Central Bank of Egypt (CBE) corpus. Two annotators from the CBE subgroup then applied the Bank of Canada (BoC) guide to these sentences, producing an initial set of labels. Independently, a separate team of annotators—using the correct CBE guide—labeled the same 500 sentences to create a gold-standard reference. Finally, we compared the BoC-guided annotations against this CBE ground truth to quantify the loss in labeling accuracy when using the wrong guide, thereby demonstrating the necessity of bank-specific annotation instructions.

As a result of this experiment, we observe that about a third of the sentences for *Stance Detection* and about 10% for *Temporal Classification* were mislabeled. Since the *Uncertainty Estimation* task is relatively similar for each bank, we observe that the number of mislabeled sentences is only 37/500. This demonstrates that the annotation guides have significant differences that cannot be overlooked. It is apparent that misunderstanding specific details such as inflation targets and policy mandates can significantly impact the annotation label, highlighting the need for unique annotation guides. A sample of mislabeled data is in Table 8.

Table 8: Sample of mislabeled sentences between groups one and two

Sentence	Task	Incorrect Label	Correct Label
annual headline urban inflation declined in october and november 2021, to record 6.3 percent and 5.6 percent from 6.6 percent in september 2021, respectively.	Stance Detection	Neutral	Dovish
meanwhile, core inflation declined to 7.9 percent in november 2018 from 8.9 percent in october 2018, recording the lowest rate since february 2016.	Stance Detection	Dovish	Hawkish
in addition to oil price developments, risks from the external economy continue to include the pace of tightening financial conditions.	Temporal Classification	Not forward looking	Forward looking
moreover, the mpc is mindful of the underlying inflationary pressures which continue to tilt the balance of risks to the upside.	Uncertainty Estimation	Uncertain	Certain

B.8.2 4-Gram Semantic Overlap

To further evaluate the degree of similarity between the annotation guides, we compute the pairwise 4-gram Jaccard similarity scores, inspired by the work of [Shah et al., 2024, Choubey et al., 2023]. After converting the annotation guides (Appendix H) to text and removing the latex jargon, we run the N-gram Jaccard semantic between pairs of all central banks’ annotation guides. As Figure 5 illustrates, there is extremely low similarity, ranging from 0.0055 to 0.1088, between the central banks’ annotation guides. Despite having the same formatting guidelines, each annotation guide has unique concepts and textual information, reinforcing the need for unique annotation guides.

B.9 Why Irrelevant Label for Stance?

Unlike other research done in the past, we utilize an alternative approach to clean the sentence level data that serves as the input for our annotation work. As mentioned in Appendix B.3, we filter the data on token length, and then use the sentence corpus as a whole rather than breaking it down

4-gram Jaccard Similarity Between Banks (Clustered)

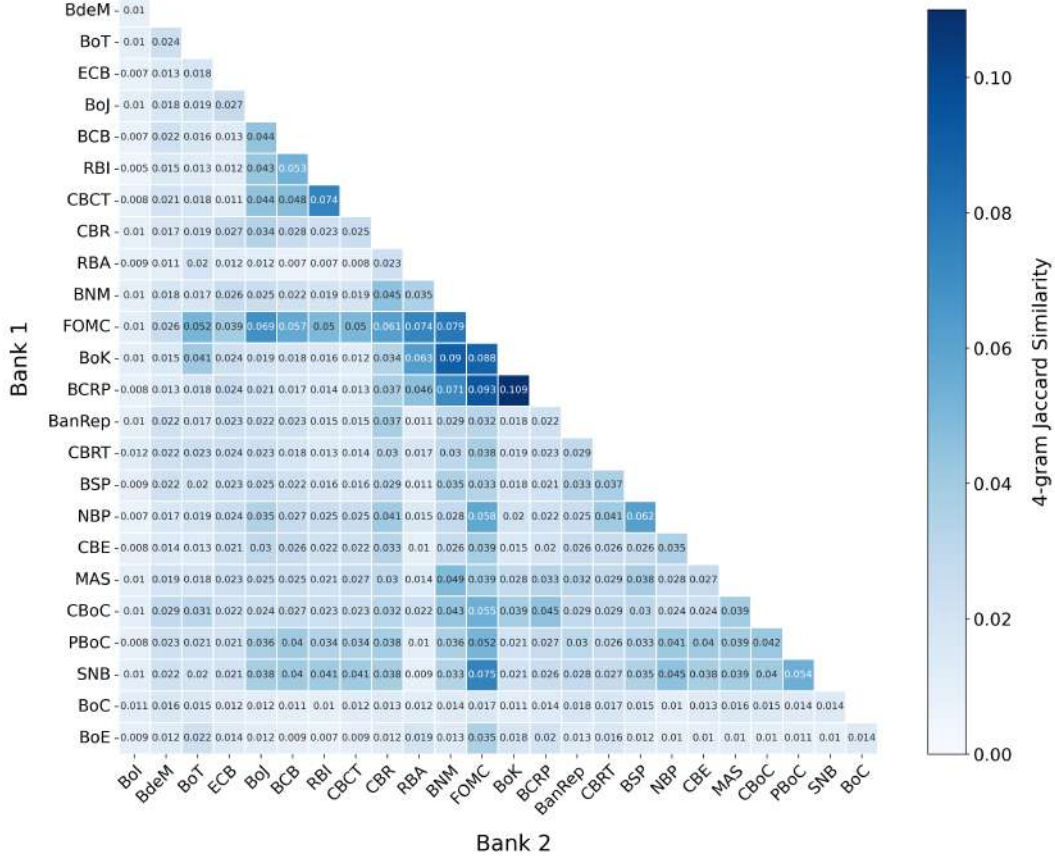


Figure 5: Pairwise 4-gram Jaccard similarity scores between annotation guides show extremely low overlap, ranging from 0.0055 to 0.1088. These results support our decision for unique annotation guides.

951 using word filtering as suggested by the work of [Gorodnichenko et al. \[2023b\]](#). Instead, we use
 952 a novel approach by adding an additional label in our annotations for Stance Detection called
 953 Irrelevant. Since we are analyzing the monetary policy communications from central banks across
 954 the world in which a specific dictionary such as [Gorodnichenko et al. \[2023b\]](#) has not previously
 955 been strictly created or defined, it is not appropriate for us to generalize the dictionary to central bank
 956 communications from around the world. Additionally, we measure the central bank communications
 957 across additional features namely forward looking nature and certainty for which word filtering is not
 958 applicable. There may be sentences that are irrelevant to monetary policy but contribute to Temporal
 959 Classification and Uncertainty Estimation tasks.

960 Moreover, we observe that using the dictionary from [Coibion et al. \[2022\]](#) rather than our classification
 961 of Irrelevant sentences as a label causes a significant number of sentences to be left out, as seen
 962 in table 9. Using the dictionary would limit the scope of our work and cause false negatives to
 963 be removed. We include a sample of sentences in Table 10 that remain in our dataset and contain
 964 valuable central bank communication.

Table 9: Comparison of number of relevant and irrelevant sentences using Word Filtering method versus the Irrelevant label for sentences pertaining to each FOMC Chair. This analysis was specifically conducted for the FOMC meeting minute data.

FED Chair — Alan Greenspan			
	Relevant	Irrelevant	Sum
Word Filter	118	122	230
Ground Truth	228	12	230
FED Chair — Ben Bernanke			
Word Filter	131	169	300
Ground Truth	289	11	300
FED Chair — Janet Yellen			
Word Filter	103	88	191
Ground Truth	183	8	191
FED Chair — Jerome Powell			
Word Filter	117	136	252
Ground Truth	245	8	252
Total			
Word Filter	469	531	1000
Ground Truth	961	39	1000

Table 10: Sample of sentences that were removed according to Yuriy’s dictionary but were included in our sampling corpus.

Sentence	Stance Detection	Temporal Classification	Uncertainty Estimation
In June, preliminary data indicated declining nominal goods exports and imports.	Neutral	Not Forward Looking	Certain
The core component of the CPI also accelerated in January and on a year-over-year basis, but by lesser amounts than did the total index.	Hawkish	Not Forward Looking	Certain
Exports of industrial supplies and consumer goods also rose strongly, while exports of services expanded modestly.	Dovish	Not Forward Looking	Certain
Real GDP was anticipated to increase at a rate noticeably below its potential in 2008.	Dovish	Forward Looking	Uncertain

965 C Models and Benchmarking

966 C.1 Prompts for Benchmarking

967 We present all the prompts used for benchmarking for all our experiments including Zero-shot
968 prompts, Few-Shot prompts, and prompts with Annotation Guides.

969 C.1.1 Zero-Shot Prompt: Stance Detection

Stance Detection Annotation Prompt

```
""You are given a sentence related to the {bank_name}'s monetary policy meeting.  
Your task is to classify its monetary policy stance and briefly justify  
your choice.
```

```
Input:
```

```
- Sentence: {sentence}
```

```
Instructions:
```

```
1. Assign one of the following labels under the 'label' key:  
"hawkish", "neutral", "dovish", or "irrelevant".
```

```
2. Provide a concise explanation for your classification using the  
'justification' key. Limit the justification to one sentence.
```

```
3. Your output must follow this structure:
```

```
{  
  "label": "hawkish | dovish | neutral | irrelevant",  
  "justification": "One-sentence explanation for the assigned label"  
}
```

Figure 6: Zero-shot prompt used for Stance Detection by a LLM across all central banks. For banks with names ending in *s*, such as the Central Bank of the Philippines, an apostrophe was added without the *s* (e.g., `bank_name'`) to ensure grammatical correctness.

```

Temporal Classification Annotation Prompt

"""You are given a sentence related to the {bank_name}'s monetary policy meeting.
Your task is to classify whether it is forward looking or not forward looking
and briefly justify your choice.

Input:
- Sentence: {sentence}

Instructions:
1. Assign one of the following labels under the 'label' key:
"forward looking", "not forward looking".

2. Provide a concise explanation for your classification using the
'justification' key. Limit the justification to one sentence.

3. Your output must follow this structure:
{
  "label": "forward looking | not forward looking",
  "justification": "One-sentence explanation for the assigned label"
}"""

```

Figure 7: Zero-shot prompt used for Temporal Classification by a LLM across all central banks. For banks with names ending in *s*, such as the Central Bank of the Philippines, an apostrophe was added without the *s* (e.g., `bank_name'`) to ensure grammatical correctness.

```
Uncertainty Estimation Annotation Prompt

"""You are given a sentence related to the {bank_name}'s monetary policy meeting.
Your task is to classify whether it is certain or uncertain
and briefly justify your choice.

Input:
- Sentence: {sentence}

Instructions:
1. Assign one of the following labels under the 'label' key:
"ertain", "uncertain".

2. Provide a concise explanation for your classification using the
'justification' key. Limit the justification to one sentence.

3. Your output must follow this structure:
{
  "label": "ertain | uncertain",
  "justification": "One-sentence explanation for the assigned label"
}"""
```

Figure 8: Zero-shot prompt used for Uncertainty Estimation by a LLM across all central banks. For banks with names ending in *s*, such as the Central Bank of the Philippines, an apostrophe was added without the *s* (e.g., bank_name') to ensure grammatical correctness.

Prompt Structure for Few-Shot prompting. This exact structure was used across all tasks.

[SYSTEM INPUT]

""You are given one sentence related to {bank_name}'s monetary policy meeting. Your task is to classify whether it is {lbl_map[feature]} and briefly justify your choice.

Instructions:

1. Assign one of the following labels under the 'label' key: {lbl_map[feature]}
2. Provide a concise explanation for your classification using the 'justification' key. Limit the justification to one sentence.
3. Your output must follow this structure:

```
{
  "label": "{instruct_map[feature]}",
  "justification": "One-sentence explanation for the assigned label"
}
```

Example:

bank_name: Bank Negara Malaysia.
label: stance

[USER INPUT]

Input:

Few-shot Examples
{examples}

Sentence
{sentence}

Example:

###Few-shot Examples

Sentence: the ceiling and floor rates of the corridor of the opr are correspondingly reduced to 2.75 percent and 2.25 percent, respectively
Label: neutral

Sentence: since march 2020, bank negara malaysia has provided additional liquidity of approximately rm42 billion into the domestic financial markets, via various tools including outright purchase of government securities, reverse repos and the reduction in statutory reserve requirement
Label: dovish

Sentence: the meeting also approved the schedule of mpc meetings for 2014
Label: irrelevant

Sentence: downside risks to this outlook, however, continue to persist
Label: hawkish

###Sentence

the mpc recognises that there are downside risks in the global economic and financial environment and is closely monitoring and assessing their implications on domestic price stability and growth

Figure 9: Prompting Structure for Few-Shot prompting using few-shot examples from the training set for that particular seed.

Prompt Structure for prompting with Annotation Guides. This exact structure was used across all tasks.

[SYSTEM INPUT]

You are given one sentence related to {bank_name}'s monetary policy meeting. You are also given an annotation guide for the task: {task}. Strictly follow the guide; do not invent new criteria or labels. Your task is to classify whether it is {lbl_map[feature]} and briefly justify your choice.

Instructions:

1. Assign one of the following labels under the 'label' key: {lbl_map[feature]}
2. Provide a concise explanation for your classification using the 'justification' key. Limit the justification to one sentence.
3. Your output must follow this structure:

```
{{
"label": "{instruct_map[feature]}",
"justification": "One-sentence explanation for the assigned label"
}}
```

Example:

bank_name: Reserve Bank of India.
label: stance

[USER INPUT]

Input:

Annotation Guide: {guide_text}

Sentence

{sentence}

Example:

###Annotation Guide

\mptext{rbi}{fourteen} Inflation,
Repo Rate, Reverse Repo Rate, Cash Reserve Ratio, Statutory Liquidity Ratio,
GDP Growth Forecast, Monetary Policy Measures,
Export Performance, Manufacturing Activity, Consumer Demand,
Employment Levels, Commodity Prices,
Credit Growth, Exchange Rate:

\begin{itemize}

\item \emph{Inflation}: A sentence pertaining to the general increase
in prices for goods and services; the decrease in the purchasing power
of a currency...

















###Sentence

the outlook for aggregate demand is progressively improving but the slack is
large: output is still below pre-covid level and the recovery is uneven and
critically dependent upon policy support.

Figure 10: Prompting Structure for prompting with Annotation Guides using the annotation guide for that particular central bank and task.

975 Details about each model such as model type, size, abbreviations used in results' tables, quantization
976 settings for LLMs, context windows (in tokens) for LLMs, and cutoff dates for LLMs are provided in
977 Table 11.

Table 11: Comparison of fine-tuned models used to classify sentences across notable characteristics. *Type* is the type of model, whether base or large, or closed or open source. *Size* is number of parameters in the model. *Abbreviation* is the abbreviated name of the model used in our paper. *Note* specifies if a model incorporates reasoning model or a mixture of experts (MOE) architectures. * means that the date for the knowledge cutoff was unable to be found. Since FinMA is fine-tuned on Llama-7B, we use the *cutoff date* for Llama-7B, despite FinMA undergoing extensive further training. The horizontal separation is to separate the PLMs (top) and the LLMs (bottom) for further analysis on the specific LLMs used.

Model	Type	Size	Abbreviation	Quantization	Context Window(Tokens)	Note	Cutoff
 ModernBERT-base	Base	150m	MBB				
 bert-base-uncased	Base	110m	BB				
 finbert-pretrain	Base	110m	FB				
 roberta-base	Base	125m	RBB				
 ModernBERT-large	Large	396m	MBL				
 bert-large	Large	340m	BL				
 roberta-large	Large	355m	RBL				
 finma-7b-full	Closed Source	-	FM	-	-	-	3/*/2023
 gemini-2.0-flash	Closed Source	-	Gem	-	1M	-	8/*/2024
 gpt-4o-2024-08-06	Closed Source	-	4o	-	128k	-	9/30/2023
 gpt-4.1-2025-04-14	Closed Source	-	4.1	-	1M	-	5/31/2024
 gpt-4.1-mini-2025-04-14	Closed Source	-	4.1M	-	1M	-	5/31/2024
 DeepSeek-V3(DeepSeek-V3-0324)	Open Source	671b	DS	FP8	131k	MoE	7/1/2024
 Qwen2.5-72B-Instruct-Turbo	Open Source	72.7b	Qwen	FP8	32.768k	-	-
 Llama-3-70b-chat-hf	Open Source	70b	L3	FP16	8k	-	12/*/2023
 Llama-4-Scout-17B-16E-Instruct	Open Source	405b	L4S	FP16	1M	-	8/*/2024

C.3 Model Specifications, Compute, and Providers

We use a 700–150–150 train–validation–test split for all PLM experiments. All models are trained using cross-entropy loss as the criteria and optimized with AdamW [Loshchilov and Hutter, 2019]. Training runs for a maximum of 50 epochs with early stopping based on validation performance. The best-performing weights and results are saved locally when running experiments for each hyperparameter combination.

Pretrained Language Models (PLMs). All PLM experiments are conducted using PyTorch [Paszke et al., 2019] on NVIDIA H200 GPUs. Models are initialized from their publicly available checkpoints via the Hugging Face Transformers library [Wolf et al., 2020]. We perform a grid search over seeds, batch sizes, and learning rates:

- seeds = [5768, 78516, 944601]
- batch sizes = [16, 32]
- learning rates = [1e-5, 1e-6]

Large Language Models (LLMs). For LLMs, we use LiteLLM [BerriAI, 2024] as a unified API router. Inference is conducted with the following hyperparameters unless otherwise noted:

- max tokens = 128.
- temperature = 0.0
- batch size = 1
- Other parameters such as top-p, top-k, and repetition penalty are kept at model-specific defaults.

Note: Temperature is fixed to 0.0 across all models. For proprietary models such as ChatGPT, actual behavior may vary slightly due to backend constraints.

LLM Providers.

- **Together.ai API:** Used for open-source models including the Llama family, DeepSeek, and Qwen2.5.

- 1003 • **OpenAI API:** Used for GPT-4o, GPT-4.1, and GPT-4.1-Mini.
- 1004 • **Gemini SDK:** Used for Gemini 2.0 Flash; accessed via the official Google Gemini
- 1005 Python client, not routed through LiteLLM.

1006 C.4 Results

1007 We present the results of the benchmarking experiments for each task (Stance Detection,
1008 Temporal Classification, and Uncertainty Estimation) for both General and Bank-
1009 Specific Setups. The results for the Stance Detection under the General Setup have already
1010 been displayed in Section 4.2.

1011 C.4.1 General (All-Banks) Setup Models

1012 Temporal Classification Label

1013 **General Setup – Temporal Classification Analysis.** In the General (All-Banks) Setup evaluation,
1014 Table 12 shows F1-Score (\pm std) for the Temporal Classification across 25 central banks,
1015 comparing four base PLMs, three large PLMs, four closed-source LLMs, and five open-source LLMs.
1016 On average, RoBERTa-Base ($\overline{F1} = 0.868$) is the best-performing PLM while Gemini-2.0-Flash
1017 ($\overline{F1} = 0.864$) is the best-performing LLM. Although RoBERTa-Base slightly outperforms Gemini
1018 across most metrics, the PLM-LLM gap remains small (<0.005 F1 on average), suggesting that both
1019 families can tackle this binary task equally well when fine-tuned on pooled data.

Table 12: F1-Scores across the banks for the Temporal Classification task with the standard deviation in the brackets. These results are based on the testing of a singular model on each bank’s data. The best-performing PLM and LLM in each row is highlighted in blue and green, respectively. The best-performing overall model is bolded. The model names are abbreviated as described in Table 11, Appendix C.2.

Bank	Base				Large			Closed-Source				Open-Source				
	MBB	BB	FB	RBB	MBL	BL	RBL	Gem	4o	4.1m	4.1	DS	Qwen	FM	L3	L4S
BCB	.857 (.019)	.859 (.017)	.844 (.016)	.870 (.024)	.884 (.039)	.870 (.032)	.851 (.022)	.835 (.031)	.840 (.025)	.809 (.027)	.822 (.021)	.845 (.014)	.830 (.027)	.182 (.010)	.821 (.005)	.826 (.014)
BCRP	.858 (.034)	.878 (.028)	.869 (.037)	.878 (.026)	.867 (.033)	.878 (.024)	.884 (.010)	.847 (.022)	.861 (.016)	.855 (.032)	.869 (.034)	.851 (.037)	.838 (.020)	.113 (.014)	.854 (.040)	.848 (.036)
BNM	.885 (.005)	.876 (.010)	.889 (.011)	.903 (.019)	.886 (.016)	.870 (.007)	.892 (.017)	.856 (.018)	.877 (.016)	.868 (.017)	.889 (.021)	.880 (.020)	.872 (.028)	.553 (.013)	.888 (.025)	.879 (.025)
BSP	.921 (.030)	.907 (.022)	.907 (.023)	.919 (.020)	.906 (.009)	.900 (.011)	.901 (.015)	.886 (.012)	.916 (.013)	.894 (.020)	.912 (.019)	.911 (.019)	.916 (.015)	.166 (.016)	.905 (.011)	.901 (.011)
BanRep	.885 (.025)	.884 (.016)	.866 (.031)	.883 (.027)	.909 (.025)	.870 (.017)	.885 (.016)	.845 (.054)	.890 (.045)	.884 (.020)	.885 (.042)	.875 (.026)	.881 (.025)	.148 (.024)	.877 (.029)	.867 (.036)
BoC	.876 (.031)	.864 (.031)	.868 (.036)	.890 (.039)	.875 (.034)	.891 (.027)	.882 (.040)	.852 (.051)	.882 (.035)	.830 (.030)	.871 (.018)	.861 (.037)	.876 (.036)	.353 (.023)	.858 (.036)	.848 (.028)
BoE	.782 (.018)	.787 (.024)	.768 (.018)	.780 (.019)	.782 (.006)	.765 (.022)	.778 (.012)	.715 (.018)	.761 (.008)	.771 (.022)	.793 (.014)	.763 (.019)	.777 (.003)	.142 (.030)	.765 (.017)	.766 (.006)
Bol	.854 (.022)	.874 (.041)	.878 (.007)	.863 (.015)	.873 (.028)	.866 (.030)	.865 (.036)	.829 (.026)	.875 (.028)	.896 (.025)	.879 (.014)	.884 (.036)	.858 (.021)	.260 (.024)	.851 (.018)	.875 (.027)
Bol	.788 (.054)	.779 (.025)	.792 (.025)	.790 (.025)	.792 (.026)	.783 (.024)	.786 (.029)	.753 (.008)	.786 (.022)	.775 (.037)	.778 (.017)	.778 (.017)	.789 (.037)	.322 (.017)	.763 (.031)	.764 (.020)
BoK	.882 (.007)	.875 (.014)	.885 (.017)	.896 (.005)	.888 (.022)	.882 (.009)	.891 (.008)	.872 (.005)	.900 (.009)	.885 (.030)	.900 (.009)	.908 (.012)	.904 (.008)	.369 (.020)	.891 (.008)	.853 (.014)
BdeM	.864 (.023)	.859 (.009)	.868 (.024)	.852 (.032)	.871 (.022)	.853 (.025)	.875 (.011)	.809 (.016)	.875 (.018)	.862 (.033)	.858 (.027)	.840 (.038)	.862 (.028)	.307 (.019)	.862 (.028)	.830 (.026)
BoT	.770 (.056)	.762 (.043)	.743 (.054)	.753 (.048)	.763 (.046)	.759 (.030)	.750 (.053)	.738 (.025)	.748 (.029)	.722 (.025)	.752 (.025)	.732 (.035)	.751 (.018)	.427 (.007)	.727 (.033)	.735 (.031)
CBCT	.833 (.029)	.829 (.039)	.831 (.029)	.824 (.039)	.806 (.030)	.823 (.039)	.828 (.029)	.831 (.028)	.831 (.019)	.810 (.027)	.820 (.019)	.822 (.037)	.818 (.035)	.306 (.025)	.825 (.021)	.809 (.015)
CBE	.926 (.021)	.913 (.020)	.925 (.019)	.914 (.017)	.923 (.025)	.904 (.024)	.929 (.028)	.913 (.011)	.935 (.006)	.943 (.011)	.933 (.009)	.943 (.014)	.928 (.007)	.182 (.025)	.935 (.008)	.936 (.006)
CBR	.887 (.020)	.887 (.018)	.880 (.032)	.876 (.018)	.887 (.020)	.904 (.023)	.885 (.030)	.857 (.040)	.851 (.008)	.838 (.029)	.849 (.019)	.840 (.016)	.857 (.031)	.235 (.027)	.851 (.027)	.861 (.024)
CBRT	.926 (.007)	.946 (.008)	.927 (.024)	.947 (.005)	.923 (.009)	.946 (.006)	.953 (.008)	.919 (.006)	.936 (.008)	.901 (.015)	.922 (.027)	.919 (.010)	.931 (.006)	.132 (.024)	.935 (.009)	.915 (.015)
CBoC	.811 (.014)	.812 (.011)	.808 (.014)	.820 (.027)	.818 (.015)	.789 (.025)	.802 (.028)	.798 (.009)	.835 (.007)	.793 (.021)	.832 (.008)	.813 (.020)	.809 (.027)	.214 (.017)	.797 (.007)	.794 (.011)
ECB	.795 (.010)	.783 (.018)	.774 (.010)	.811 (.011)	.790 (.024)	.791 (.008)	.774 (.009)	.741 (.045)	.781 (.022)	.763 (.023)	.784 (.014)	.771 (.013)	.779 (.023)	.267 (.051)	.788 (.019)	.780 (.011)
FOMC	.881 (.017)	.904 (.005)	.888 (.027)	.899 (.019)	.887 (.024)	.890 (.007)	.890 (.023)	.844 (.022)	.907 (.019)	.865 (.022)	.904 (.017)	.890 (.017)	.909 (.025)	.184 (.010)	.882 (.019)	.883 (.008)
MAS	.936 (.003)	.942 (.023)	.938 (.017)	.933 (.018)	.936 (.013)	.932 (.009)	.928 (.011)	.903 (.015)	.927 (.020)	.926 (.014)	.931 (.017)	.933 (.025)	.951 (.017)	.371 (.050)	.920 (.009)	.900 (.011)
NBP	.887 (.011)	.901 (.015)	.879 (.028)	.894 (.007)	.899 (.010)	.887 (.013)	.887 (.006)	.868 (.009)	.880 (.016)	.861 (.011)	.913 (.005)	.877 (.006)	.900 (.014)	.395 (.019)	.896 (.019)	.876 (.011)
PBoC	.914 (.011)	.913 (.008)	.907 (.006)	.919 (.012)	.904 (.019)	.913 (.019)	.923 (.015)	.863 (.017)	.914 (.008)	.900 (.013)	.908 (.021)	.903 (.026)	.917 (.015)	.580 (.036)	.900 (.016)	.879 (.007)
RBA	.913 (.028)	.935 (.022)	.909 (.004)	.920 (.027)	.924 (.013)	.918 (.035)	.930 (.033)	.891 (.021)	.935 (.009)	.917 (.017)	.929 (.008)	.921 (.009)	.920 (.015)	.183 (.032)	.924 (.008)	.912 (.018)
RBI	.812 (.005)	.828 (.013)	.798 (.018)	.817 (.011)	.826 (.019)	.813 (.016)	.835 (.021)	.784 (.020)	.802 (.006)	.795 (.011)	.807 (.019)	.807 (.022)	.800 (.026)	.389 (.010)	.807 (.010)	.790 (.012)
SNB	.837 (.020)	.843 (.013)	.850 (.014)	.848 (.010)	.836 (.013)	.827 (.010)	.838 (.011)	.800 (.018)	.853 (.019)	.854 (.009)	.850 (.023)	.859 (.013)	.853 (.006)	.425 (.047)	.860 (.011)	.822 (.026)
Average	.863 (.049)	.866 (.052)	.860 (.053)	.868 (.051)	.866 (.050)	.861 (.053)	.865 (.054)	.834 (.054)	.864 (.054)	.849 (.056)	.863 (.052)	.857 (.055)	.861 (.055)	.288 (.126)	0.855 (.055)	.846 (.052)

1020 Uncertainty Estimation Task

1021 **General Setup – Uncertainty Estimation Analysis.** Table 13 reports F1-Score (\pm std) on the
 1022 Uncertainty Estimation for four base PLMs, three large PLMs, four closed-source LLMs, and
 1023 five open-source LLMs, tested on pooled all-banks data. RoBERTa-Large ($F1 = 0.846$) again
 1024 ranks highest among PLMs, with Llama-3-70B-Chat ($F1 = 0.816$) performing best among LLMs.
 1025 The overall F1-Scores for this task are higher than those for Stance Detection, but the relative
 1026 ordering across models resembles the Stance Detection task: PLMs outperform LLMs on average.
 1027 Uncertainty Estimation proves to be more difficult than Temporal Classification due to
 1028 lower average scores across all models.

Table 13: F1-Scores across the banks for the Uncertainty Estimation task with the standard deviation in the brackets. These results are based on the testing of a singular model on each bank’s data. The best-performing PLM and LLM in each row is highlighted in blue and green, respectively. The best-performing overall model is bolded. The model names are abbreviated as described in Table 11, Appendix C.2.

Bank	Base				Large			Closed-Source				Open-Source				
	MBB	BB	FB	RBB	MBL	BL	RBL	Gem	4o	4.1m	4.1	DS	Qwen	FM	L3	L4S
BCB	.863 (.002)	.869 (.006)	.879 (.005)	.890 (.006)	.888 (.023)	.894 (.031)	.873 (.005)	.741 (.025)	.607 (.035)	.779 (.020)	.749 (.006)	.692 (.004)	.670 (.014)	.696 (.029)	.819 (.010)	.766 (.019)
BCRP	.912 (.032)	.902 (.030)	.916 (.034)	.913 (.037)	.918 (.031)	.916 (.032)	.903 (.032)	.852 (.011)	.816 (.030)	.857 (.025)	.847 (.024)	.838 (.034)	.807 (.027)	.884 (.026)	.898 (.016)	.888 (.025)
BNM	.813 (.036)	.813 (.029)	.805 (.028)	.812 (.026)	.808 (.029)	.799 (.046)	.812 (.020)	.693 (.006)	.676 (.002)	.769 (.009)	.793 (.010)	.760 (.015)	.616 (.045)	.660 (.021)	.784 (.010)	.808 (.013)
BSP	.889 (.019)	.893 (.010)	.893 (.032)	.879 (.020)	.881 (.017)	.894 (.010)	.876 (.010)	.822 (.033)	.690 (.033)	.789 (.024)	.779 (.028)	.788 (.024)	.753 (.030)	.785 (.024)	.877 (.024)	.828 (.007)
BanRep	.868 (.013)	.862 (.010)	.850 (.023)	.872 (.018)	.882 (.007)	.864 (.006)	.880 (.011)	.822 (.019)	.737 (.040)	.843 (.005)	.821 (.031)	.798 (.022)	.764 (.031)	.718 (.019)	.855 (.030)	.782 (.044)
BoC	.823 (.048)	.837 (.015)	.842 (.033)	.857 (.027)	.821 (.039)	.855 (.028)	.858 (.014)	.746 (.010)	.725 (.022)	.825 (.031)	.799 (.035)	.737 (.014)	.702 (.010)	.506 (.041)	.751 (.011)	.699 (.035)
BoE	.692 (.019)	.710 (.036)	.696 (.033)	.676 (.031)	.700 (.028)	.713 (.024)	.721 (.025)	.669 (.035)	.663 (.019)	.681 (.043)	.687 (.043)	.662 (.025)	.698 (.028)	.595 (.036)	.730 (.022)	.702 (.013)
Bol	.712 (.010)	.729 (.015)	.732 (.016)	.733 (.005)	.705 (.013)	.705 (.021)	.732 (.014)	.665 (.017)	.643 (.015)	.709 (.026)	.685 (.010)	.666 (.015)	.651 (.001)	.660 (.002)	.724 (.005)	.673 (.015)
BoJ	.823 (.022)	.835 (.006)	.840 (.005)	.825 (.010)	.828 (.004)	.836 (.005)	.821 (.018)	.621 (.025)	.553 (.028)	.625 (.042)	.652 (.029)	.512 (.045)	.604 (.048)	.670 (.015)	.782 (.009)	.718 (.021)
BoK	.856 (.031)	.860 (.031)	.861 (.021)	.866 (.028)	.865 (.014)	.846 (.018)	.879 (.033)	.658 (.026)	.561 (.054)	.650 (.030)	.721 (.031)	.582 (.042)	.637 (.027)	.686 (.020)	.801 (.008)	.760 (.019)
BdeM	.736 (.036)	.713 (.044)	.722 (.049)	.738 (.039)	.713 (.053)	.723 (.040)	.719 (.056)	.666 (.055)	.541 (.027)	.685 (.023)	.703 (.035)	.454 (.007)	.575 (.040)	.629 (.067)	.750 (.012)	.639 (.051)
BoT	.657 (.025)	.662 (.011)	.662 (.025)	.665 (.024)	.683 (.009)	.661 (.012)	.681 (.005)	.590 (.021)	.541 (.045)	.571 (.028)	.600 (.023)	.571 (.043)	.564 (.036)	.623 (.024)	.662 (.011)	.630 (.031)
CBCT	.800 (.001)	.821 (.031)	.823 (.035)	.802 (.022)	.791 (.031)	.816 (.012)	.810 (.019)	.780 (.031)	.719 (.033)	.785 (.036)	.800 (.026)	.792 (.026)	.757 (.025)	.652 (.020)	.784 (.039)	.756 (.011)
CBE	.920 (.002)	.900 (.015)	.919 (.023)	.906 (.012)	.908 (.012)	.911 (.021)	.917 (.031)	.829 (.031)	.815 (.014)	.862 (.020)	.852 (.024)	.839 (.015)	.838 (.026)	.806 (.017)	.904 (.027)	.874 (.021)
CBR	.910 (.032)	.913 (.034)	.895 (.027)	.908 (.023)	.899 (.017)	.916 (.025)	.902 (.022)	.701 (.043)	.662 (.029)	.778 (.030)	.767 (.055)	.807 (.054)	.713 (.037)	.817 (.014)	.830 (.041)	.797 (.024)
CBRT	.927 (.018)	.935 (.028)	.936 (.009)	.935 (.008)	.932 (.016)	.928 (.009)	.951 (.010)	.851 (.021)	.780 (.011)	.884 (.023)	.884 (.014)	.843 (.007)	.843 (.010)	.792 (.009)	.906 (.019)	.870 (.014)
CBoC	.862 (.022)	.859 (.036)	.857 (.025)	.866 (.021)	.864 (.038)	.854 (.012)	.855 (.017)	.754 (.041)	.577 (.048)	.764 (.045)	.741 (.038)	.623 (.068)	.682 (.060)	.713 (.019)	.816 (.028)	.705 (.042)
ECB	.885 (.027)	.863 (.031)	.878 (.023)	.876 (.029)	.858 (.025)	.861 (.058)	.881 (.036)	.735 (.019)	.645 (.006)	.722 (.027)	.762 (.023)	.669 (.039)	.735 (.010)	.701 (.036)	.841 (.007)	.787 (.013)
FOMC	.808 (.020)	.838 (.012)	.791 (.025)	.809 (.020)	.818 (.010)	.792 (.031)	.835 (.019)	.818 (.019)	.763 (.045)	.806 (.016)	.779 (.034)	.718 (.048)	.818 (.032)	.671 (.033)	.839 (.022)	.825 (.011)
MAS	.827 (.029)	.843 (.013)	.847 (.016)	.834 (.023)	.826 (.042)	.837 (.019)	.840 (.027)	.798 (.043)	.803 (.027)	.822 (.036)	.828 (.035)	.809 (.024)	.800 (.029)	.602 (.039)	.858 (.035)	.857 (.031)
NBP	.840 (.016)	.850 (.016)	.838 (.015)	.856 (.006)	.846 (.015)	.866 (.005)	.851 (.011)	.760 (.013)	.610 (.021)	.725 (.007)	.745 (.023)	.623 (.026)	.689 (.033)	.607 (.023)	.813 (.016)	.774 (.019)
PBoC	.944 (.008)	.948 (.011)	.945 (.012)	.940 (.004)	.943 (.007)	.945 (.006)	.950 (.007)	.641 (.046)	.749 (.017)	.851 (.013)	.801 (.023)	.869 (.022)	.844 (.023)	.927 (.021)	.955 (.010)	.942 (.012)
RBA	.867 (.057)	.882 (.054)	.834 (.050)	.894 (.043)	.842 (.068)	.865 (.054)	.890 (.021)	.847 (.015)	.782 (.021)	.851 (.031)	.852 (.032)	.778 (.019)	.824 (.029)	.671 (.022)	.854 (.039)	.847 (.019)
RBI	.862 (.028)	.835 (.027)	.850 (.023)	.834 (.015)	.851 (.025)	.833 (.018)	.851 (.023)	.662 (.018)	.584 (.017)	.700 (.024)	.686 (.021)	.616 (.020)	.599 (.042)	.750 (.030)	.755 (.023)	.702 (.029)
SNB	.853 (.014)	.868 (.015)	.863 (.011)	.869 (.015)	.855 (.019)	.852 (.014)	.864 (.017)	.703 (.017)	.680 (.001)	.779 (.007)	.759 (.010)	.721 (.015)	.651 (.010)	.733 (.008)	.805 (.006)	.779 (.025)
Average	.838 (.073)	.842 (.071)	.839 (.072)	.842 (.073)	.837 (.072)	.839 (.073)	.846 (.070)	.737 (.077)	.677 (.088)	.764 (.079)	.764 (.067)	.711 (.109)	.713 (.087)	.702 (.093)	.816 (.066)	.776 (.078)

1029 C.4.2 Specific Bank Setup Models

1030 Stance Label

1031 **Bank-Specific Setup – Stance Detection.** Table 14 shows bank-by-bank F1-Score (\pm std)
 1032 for Hawkish/Dovish/Neutral/Irrelevant, comparing four base PLMs, three large PLMs,
 1033 four closed-source LLMs, and five open-source LLMs, each fine-tuned on that bank’s data.
 1034 RoBERTa-Large ($\overline{F1} = 0.694$) remains the best-performing PLM, and Llama-3-70B-Chat
 1035 ($\overline{F1} = 0.620$) leads the LLMs, consistent with the results from the General Setup. The relative
 1036 ranking of models is largely preserved, with PLMs outperforming LLMs overall. While the
 1037 performance gap between PLMs and LLMs narrows slightly (mean $\Delta = +0.074$) in this setting, it
 1038 remains clear that PLMs benefit more from fine-tuning on the aggregated dataset. It is important to
 1039 note that LLMs are evaluated using the same zero-shot prompts in both the General and Bank-Specific
 1040 Setups.

Table 14: F1-Scores of models classifying sentences from each central bank with the monetary policy stance label of Hawkish/Dovish/Neutral/Irrelevant with standard deviation in the parenthesis. The best-performing PLM and LLM in each row is highlighted in blue and green, respectively. The best-performing overall model is bolded. The model names are abbreviated as described in Table 11, Appendix C.2.

Bank	Base				Large			Closed-Source				Open-Source				
	MBB	BB	FB	RBB	MBL	BL	RBL	Gem	4o	4.1m	4.1	DS	Qwen	FM	L3	L4S
BCB	.525 (.023)	.456 (.023)	.518 (.067)	.572 (.042)	.507 (.055)	.527 (.064)	.623 (.012)	.528 (.027)	.498 (.017)	.462 (.020)	.504 (.017)	.613 (.021)	.525 (.028)	.350 (.034)	.503 (.021)	.589 (.049)
BCRP	.721 (.035)	.769 (.029)	.768 (.033)	.784 (.032)	.729 (.089)	.741 (.030)	.809 (.032)	.675 (.004)	.628 (.008)	.634 (.035)	.665 (.004)	.666 (.010)	.503 (.062)	.301 (.031)	.641 (.014)	.620 (.043)
BNM	.590 (.046)	.632 (.076)	.600 (.062)	.630 (.023)	.667 (.026)	.607 (.078)	.642 (.058)	.409 (.006)	.443 (.027)	.430 (.007)	.409 (.025)	.475 (.013)	.333 (.026)	.160 (.033)	.567 (.025)	.435 (.005)
BSP	.521 (.021)	.608 (.076)	.667 (.073)	.677 (.043)	.540 (.028)	.560 (.073)	.698 (.048)	.424 (.042)	.420 (.069)	.451 (.039)	.514 (.076)	.534 (.029)	.380 (.015)	.245 (.027)	.584 (.042)	.500 (.035)
BanRep	.472 (.009)	.585 (.031)	.644 (.025)	.637 (.088)	.515 (.058)	.558 (.042)	.691 (.037)	.515 (.021)	.455 (.031)	.520 (.036)	.553 (.015)	.570 (.037)	.450 (.023)	.230 (.033)	.573 (.038)	.423 (.033)
BoC	.582 (.073)	.633 (.091)	.653 (.021)	.693 (.034)	.635 (.050)	.708 (.022)	.728 (.009)	.629 (.069)	.647 (.052)	.641 (.052)	.657 (.028)	.657 (.026)	.524 (.038)	.264 (.029)	.669 (.043)	.644 (.024)
BoE	.560 (.033)	.595 (.054)	.646 (.049)	.677 (.049)	.659 (.041)	.653 (.018)	.702 (.041)	.543 (.026)	.524 (.031)	.537 (.048)	.602 (.031)	.543 (.070)	.396 (.021)	.129 (.026)	.661 (.044)	.518 (.045)
Bol	.460 (.082)	.543 (.050)	.556 (.062)	.587 (.019)	.539 (.071)	.570 (.044)	.635 (.048)	.474 (.011)	.460 (.005)	.433 (.032)	.526 (.013)	.482 (.001)	.329 (.025)	.085 (.011)	.594 (.023)	.430 (.008)
BoJ	.495 (.009)	.510 (.069)	.602 (.040)	.608 (.017)	.544 (.036)	.515 (.034)	.606 (.032)	.524 (.010)	.545 (.021)	.465 (.028)	.565 (.008)	.498 (.009)	.406 (.040)	.157 (.033)	.574 (.027)	.507 (.026)
BoK	.556 (.048)	.572 (.077)	.548 (.056)	.645 (.012)	.581 (.026)	.610 (.030)	.704 (.052)	.646 (.040)	.648 (.030)	.594 (.066)	.678 (.016)	.629 (.016)	.466 (.076)	.181 (.031)	.632 (.032)	.592 (.047)
BdeM	.517 (.026)	.437 (.082)	.596 (.009)	.618 (.051)	.559 (.031)	.536 (.022)	.684 (.059)	.596 (.009)	.602 (.024)	.509 (.023)	.626 (.016)	.669 (.034)	.447 (.047)	.118 (.018)	.642 (.012)	.552 (.013)
BoT	.597 (.010)	.604 (.033)	.666 (.025)	.654 (.118)	.636 (.045)	.599 (.014)	.680 (.065)	.547 (.004)	.549 (.032)	.551 (.039)	.573 (.012)	.581 (.029)	.484 (.038)	.258 (.026)	.596 (.009)	.577 (.030)
CBCT	.488 (.067)	.564 (.014)	.613 (.027)	.631 (.037)	.564 (.054)	.557 (.038)	.644 (.032)	.451 (.044)	.474 (.044)	.474 (.031)	.485 (.026)	.522 (.022)	.388 (.049)	.180 (.037)	.556 (.015)	.475 (.051)
CBE	.708 (.052)	.676 (.029)	.729 (.076)	.749 (.014)	.733 (.035)	.736 (.062)	.783 (.028)	.629 (.037)	.672 (.036)	.581 (.045)	.648 (.036)	.636 (.007)	.352 (.056)	.142 (.014)	.702 (.021)	.594 (.024)
CBR	.675 (.085)	.652 (.042)	.666 (.059)	.750 (.035)	.671 (.085)	.716 (.027)	.796 (.042)	.759 (.015)	.749 (.027)	.693 (.026)	.701 (.049)	.794 (.028)	.573 (.035)	.146 (.022)	.772 (.029)	.665 (.013)
CBRT	.505 (.010)	.600 (.018)	.665 (.067)	.673 (.015)	.596 (.034)	.650 (.044)	.716 (.048)	.495 (.006)	.421 (.014)	.424 (.018)	.475 (.015)	.539 (.030)	.277 (.006)	.133 (.020)	.653 (.036)	.416 (.032)
CBoC	.485 (.106)	.574 (.041)	.606 (.069)	.653 (.019)	.569 (.041)	.603 (.082)	.721 (.085)	.668 (.032)	.604 (.027)	.605 (.033)	.678 (.057)	.676 (.038)	.559 (.072)	.223 (.040)	.685 (.019)	.539 (.071)
ECB	.479 (.028)	.491 (.036)	.530 (.027)	.601 (.013)	.533 (.119)	.518 (.040)	.654 (.006)	.638 (.023)	.599 (.019)	.610 (.038)	.660 (.021)	.637 (.005)	.548 (.017)	.206 (.052)	.613 (.020)	.595 (.016)
FOMC	.455 (.008)	.505 (.022)	.559 (.046)	.589 (.030)	.558 (.067)	.563 (.012)	.691 (.011)	.572 (.021)	.584 (.025)	.564 (.018)	.649 (.023)	.653 (.023)	.512 (.023)	.170 (.025)	.599 (.012)	.498 (.015)
MAS	.492 (.009)	.598 (.065)	.632 (.072)	.704 (.010)	.601 (.035)	.589 (.003)	.690 (.034)	.553 (.046)	.581 (.026)	.588 (.034)	.569 (.015)	.689 (.041)	.540 (.026)	.347 (.024)	.638 (.035)	.646 (.023)
NBP	.586 (.015)	.540 (.017)	.594 (.014)	.666 (.026)	.555 (.060)	.617 (.027)	.674 (.045)	.637 (.015)	.631 (.043)	.614 (.063)	.665 (.031)	.660 (.002)	.508 (.028)	.118 (.017)	.618 (.035)	.597 (.015)
PBoC	.727 (.025)	.742 (.030)	.736 (.023)	.762 (.025)	.744 (.019)	.730 (.014)	.764 (.030)	.492 (.046)	.559 (.037)	.531 (.026)	.535 (.033)	.592 (.037)	.379 (.017)	.128 (.018)	.613 (.033)	.446 (.024)
RBA	.531 (.072)	.516 (.041)	.586 (.059)	.685 (.023)	.549 (.024)	.553 (.022)	.642 (.038)	.531 (.049)	.478 (.079)	.483 (.058)	.553 (.074)	.537 (.055)	.358 (.049)	.133 (.020)	.614 (.034)	.495 (.057)
RBI	.519 (.050)	.593 (.015)	.584 (.043)	.640 (.056)	.608 (.067)	.590 (.038)	.673 (.043)	.489 (.025)	.519 (.041)	.509 (.026)	.495 (.027)	.542 (.016)	.431 (.008)	.231 (.058)	.581 (.030)	.557 (.032)
SNB	.573 (.026)	.612 (.046)	.680 (.023)	.653 (.022)	.530 (.056)	.674 (.013)	.704 (.025)	.635 (.003)	.601 (.015)	.640 (.037)	.652 (.016)	.643 (.024)	.554 (.029)	.252 (.008)	.612 (.014)	.607 (.051)
Average	.553 (.081)	.584 (.077)	.636 (.064)	.662 (.056)	.597 (.070)	.611 (.071)	.694 (.052)	.562 (.085)	.556 (.084)	.542 (.075)	.586 (.078)	.601 (.075)	.449 (.083)	.196 (.071)	.620 (.053)	.541 (.074)

1041 Temporal Classification Label

1042 **Bank-Specific Setup – Temporal Classification Analysis.** Table 15 reports per-bank F1-Score
 1043 (\pm std) for the Temporal Classification across four base PLMs, three large PLMs, four
 1044 closed-source LLMs, and five open-source LLMs—each fine-tuned solely on that bank’s data.
 1045 BERT-Large ($F1 = 0.861$) achieves the best PLM score for Temporal Classification, while
 1046 GPT-4o ($F1 = 0.864$) leads among LLMs. The average performance between PLMs and LLMs is
 1047 very similar (<0.005 F1-Score on the best performing models).

Table 15: F1-Scores of models classifying sentences from each central bank as Forward Looking/Not Forward Looking, with standard deviation in the parenthesis. The best-performing PLM and LLM in each row is highlighted in blue and green, respectively. The best-performing overall model is bolded. The model names are abbreviated as described in Table 11, Appendix C.2.

Bank	Base				Large			Closed-Source				Open-Source				
	MBB	BB	FB	RBB	MBL	BL	RBL	Gem	4o	4.1m	4.1	DS	Qwen	FM	L3	L4S
BCB	.818 (.027)	.821 (.060)	.863 (.015)	.843 (.050)	.840 (.021)	.866 (.027)	.869 (.023)	.835 (.031)	.840 (.025)	.809 (.027)	.822 (.021)	.845 (.014)	.830 (.027)	.182 (.010)	.821 (.005)	.826 (.014)
BCRP	.866 (.044)	.867 (.047)	.868 (.038)	.847 (.028)	.872 (.024)	.867 (.036)	.882 (.034)	.847 (.022)	.861 (.016)	.855 (.032)	.869 (.034)	.851 (.037)	.838 (.020)	.113 (.014)	.854 (.040)	.848 (.036)
BNM	.850 (.025)	.867 (.029)	.877 (.006)	.877 (.003)	.899 (.010)	.878 (.034)	.868 (.023)	.856 (.018)	.877 (.016)	.868 (.017)	.889 (.021)	.880 (.020)	.872 (.028)	.553 (.013)	.888 (.025)	.879 (.025)
BSP	.906 (.017)	.899 (.016)	.894 (.031)	.904 (.030)	.913 (.011)	.908 (.012)	.916 (.016)	.886 (.012)	.916 (.013)	.894 (.020)	.912 (.019)	.911 (.019)	.916 (.015)	.166 (.016)	.905 (.011)	.901 (.011)
BanRep	.857 (.045)	.849 (.013)	.863 (.013)	.847 (.015)	.865 (.038)	.896 (.036)	.868 (.018)	.845 (.054)	.890 (.045)	.884 (.020)	.885 (.042)	.875 (.026)	.881 (.025)	.148 (.024)	.877 (.029)	.867 (.036)
BoC	.856 (.056)	.852 (.050)	.864 (.059)	.864 (.078)	.869 (.042)	.879 (.050)	.878 (.041)	.852 (.051)	.882 (.035)	.830 (.030)	.871 (.018)	.861 (.037)	.876 (.036)	.353 (.023)	.858 (.036)	.848 (.028)
BoE	.736 (.028)	.732 (.031)	.777 (.020)	.760 (.014)	.757 (.003)	.781 (.009)	.761 (.029)	.715 (.018)	.761 (.008)	.771 (.022)	.793 (.014)	.763 (.019)	.777 (.003)	.142 (.030)	.765 (.017)	.766 (.006)
Bol	.822 (.021)	.803 (.046)	.847 (.003)	.850 (.063)	.873 (.036)	.855 (.030)	.864 (.038)	.829 (.026)	.875 (.028)	.896 (.025)	.879 (.014)	.884 (.036)	.858 (.021)	.260 (.024)	.851 (.018)	.875 (.027)
BoJ	.776 (.033)	.765 (.013)	.785 (.015)	.752 (.002)	.771 (.006)	.792 (.022)	.767 (.025)	.753 (.008)	.786 (.022)	.775 (.037)	.778 (.017)	.778 (.017)	.789 (.037)	.322 (.017)	.763 (.031)	.764 (.020)
BoK	.861 (.012)	.854 (.024)	.879 (.011)	.852 (.021)	.859 (.007)	.874 (.022)	.881 (.010)	.872 (.005)	.900 (.009)	.885 (.030)	.900 (.009)	.908 (.012)	.904 (.008)	.369 (.020)	.891 (.008)	.853 (.014)
BdeM	.801 (.032)	.794 (.026)	.829 (.044)	.822 (.058)	.821 (.016)	.864 (.024)	.837 (.035)	.809 (.016)	.875 (.018)	.862 (.033)	.858 (.027)	.840 (.038)	.862 (.028)	.307 (.019)	.862 (.028)	.830 (.026)
BoT	.712 (.062)	.700 (.034)	.734 (.085)	.719 (.040)	.719 (.061)	.734 (.086)	.742 (.067)	.738 (.025)	.748 (.029)	.722 (.025)	.752 (.025)	.732 (.035)	.751 (.018)	.427 (.007)	.727 (.033)	.735 (.031)
CBCT	.783 (.041)	.786 (.016)	.807 (.047)	.829 (.032)	.824 (.041)	.816 (.046)	.817 (.038)	.831 (.028)	.831 (.019)	.810 (.027)	.820 (.019)	.822 (.037)	.818 (.035)	.306 (.025)	.825 (.021)	.809 (.015)
CBE	.927 (.014)	.930 (.020)	.913 (.022)	.912 (.036)	.919 (.009)	.928 (.027)	.925 (.006)	.913 (.011)	.935 (.006)	.943 (.009)	.933 (.009)	.943 (.014)	.928 (.007)	.182 (.025)	.935 (.008)	.936 (.006)
CBR	.874 (.026)	.876 (.048)	.901 (.023)	.891 (.042)	.888 (.031)	.888 (.037)	.887 (.042)	.857 (.040)	.851 (.008)	.838 (.029)	.849 (.019)	.840 (.016)	.857 (.031)	.235 (.027)	.851 (.027)	.861 (.024)
CBRT	.905 (.021)	.904 (.039)	.910 (.012)	.906 (.022)	.898 (.003)	.925 (.015)	.907 (.020)	.919 (.006)	.936 (.008)	.901 (.015)	.922 (.027)	.919 (.010)	.931 (.006)	.132 (.024)	.935 (.009)	.915 (.015)
CBoC	.773 (.009)	.777 (.045)	.790 (.037)	.758 (.023)	.760 (.007)	.797 (.013)	.787 (.026)	.798 (.009)	.835 (.007)	.793 (.021)	.832 (.008)	.813 (.020)	.809 (.027)	.214 (.017)	.797 (.007)	.794 (.011)
ECB	.770 (.073)	.706 (.044)	.789 (.040)	.752 (.038)	.774 (.040)	.808 (.047)	.803 (.047)	.741 (.045)	.781 (.022)	.763 (.023)	.784 (.014)	.771 (.013)	.779 (.023)	.267 (.051)	.788 (.019)	.780 (.011)
FOMC	.850 (.035)	.863 (.040)	.886 (.027)	.866 (.020)	.875 (.040)	.866 (.022)	.895 (.033)	.844 (.022)	.907 (.019)	.865 (.022)	.904 (.017)	.890 (.017)	.909 (.025)	.184 (.010)	.882 (.019)	.883 (.008)
MAS	.920 (.013)	.929 (.020)	.944 (.011)	.944 (.022)	.960 (.015)	.942 (.023)	.938 (.017)	.903 (.015)	.927 (.020)	.926 (.014)	.931 (.017)	.933 (.025)	.951 (.017)	.371 (.050)	.920 (.009)	.900 (.011)
NBP	.861 (.004)	.873 (.011)	.886 (.032)	.863 (.018)	.871 (.020)	.875 (.034)	.908 (.028)	.868 (.009)	.880 (.016)	.861 (.011)	.913 (.005)	.877 (.006)	.900 (.014)	.395 (.019)	.896 (.019)	.876 (.011)
PBoC	.919 (.017)	.913 (.024)	.907 (.023)	.911 (.011)	.916 (.019)	.912 (.008)	.910 (.018)	.863 (.017)	.914 (.008)	.900 (.013)	.908 (.021)	.903 (.026)	.917 (.015)	.580 (.036)	.900 (.016)	.879 (.007)
RBA	.913 (.025)	.917 (.040)	.909 (.021)	.922 (.026)	.893 (.029)	.916 (.037)	.922 (.038)	.891 (.021)	.935 (.009)	.917 (.017)	.929 (.008)	.921 (.009)	.920 (.015)	.183 (.032)	.924 (.008)	.912 (.018)
RBI	.781 (.016)	.780 (.042)	.807 (.042)	.792 (.022)	.822 (.028)	.833 (.014)	.803 (.034)	.784 (.020)	.802 (.006)	.795 (.011)	.807 (.019)	.807 (.022)	.800 (.026)	.389 (.010)	.807 (.010)	.790 (.012)
SNB	.811 (.018)	.794 (.039)	.817 (.035)	.773 (.038)	.823 (.030)	.819 (.021)	.846 (.016)	.800 (.018)	.853 (.019)	.854 (.009)	.850 (.023)	.859 (.013)	.853 (.006)	.425 (.047)	.860 (.011)	.822 (.026)
Average	.838 (.061)	.834 (.068)	.854 (.053)	.842 (.062)	.851 (.059)	.861 (.052)	.859 (.056)	.834 (.054)	.864 (.054)	.849 (.056)	.863 (.052)	.857 (.055)	.861 (.055)	.288 (.126)	0.855 (.055)	.846 (.052)

Bank-Specific Setup – Uncertainty Estimation Analysis. Table 16 reports per-bank F1-Score (\pm std) on the Uncertainty Estimation task for four base PLMs, three large PLMs, four closed-source LLMs, and five open-source LLMs—each fine-tuned only on that bank’s data. The Uncertainty Estimation task exhibits the highest variance across models. Llama-3-70B-Chat ($\overline{F1} = 0.816$) is the best-performing LLM, while PLM performance is more mixed. RoBERTa-Large ($\overline{F1} = 0.820$) and FinBERT-Pretrain ($\overline{F1} = 0.820$) achieve the best single-bank result. RoBERTa-Large has a higher standard deviation (0.100) compared to FinBERT-Pretrain (0.093) but PLMs on average still outperform LLMs (mean $\Delta = +0.004$). These findings indicate that Uncertainty Estimation as a task is especially sensitive to changes in central bank phrasing, which hampers the models’ ability to generalize on this task.

Table 16: F1-Scores of models classifying sentences from each central bank as Uncertain/Certain, with standard deviation in the parenthesis. The best-performing PLM and LLM in each row is highlighted in blue and green, respectively. The best-performing overall model is bolded. The model names are abbreviated as described in Table 11, Appendix C.2.



















































Bank	Base				Large			Closed-Source				Open-Source				
	MBB	BB	FB	RBB	MBL	BL	RBL	Gem	4o	4.1m	4.1	DS	Qwen	FM	L3	L4S
 BCB	.804 (.052)	.823 (.053)	.876 (.034)	.845 (.058)	.860 (.029)	.798 (.110)	.861 (.007)	.741 (.025)	.607 (.035)	.779 (.020)	.749 (.006)	.692 (.004)	.670 (.014)	.696 (.029)	.819 (.010)	.766 (.019)
 BCRP	.910 (.032)	.908 (.036)	.912 (.033)	.915 (.035)	.923 (.027)	.920 (.016)	.916 (.038)	.852 (.011)	.816 (.030)	.857 (.025)	.847 (.024)	.838 (.034)	.807 (.027)	.884 (.026)	.898 (.016)	.888 (.025)
 BNM	.738 (.021)	.762 (.027)	.790 (.039)	.761 (.036)	.808 (.007)	.807 (.041)	.809 (.024)	.693 (.006)	.676 (.002)	.769 (.009)	.793 (.010)	.760 (.015)	.616 (.045)	.660 (.021)	.784 (.010)	.808 (.013)
 BSP	.886 (.024)	.858 (.026)	.882 (.014)	.873 (.026)	.879 (.044)	.867 (.054)	.869 (.026)	.822 (.033)	.690 (.033)	.789 (.024)	.779 (.028)	.788 (.024)	.753 (.030)	.785 (.024)	.877 (.024)	.828 (.007)
 BanRep	.792 (.041)	.845 (.013)	.848 (.021)	.832 (.019)	.843 (.024)	.850 (.012)	.868 (.020)	.822 (.019)	.737 (.040)	.843 (.005)	.821 (.031)	.798 (.022)	.764 (.031)	.718 (.019)	.855 (.030)	.782 (.044)
 BoC	.794 (.040)	.839 (.033)	.838 (.020)	.859 (.029)	.849 (.028)	.819 (.074)	.843 (.042)	.746 (.010)	.725 (.022)	.825 (.031)	.799 (.035)	.737 (.014)	.702 (.010)	.506 (.041)	.751 (.011)	.699 (.035)
 BoE	.641 (.012)	.681 (.026)	.699 (.025)	.656 (.083)	.670 (.051)	.677 (.040)	.694 (.067)	.669 (.035)	.663 (.019)	.681 (.043)	.687 (.043)	.662 (.025)	.698 (.028)	.595 (.036)	.730 (.036)	.702 (.013)
 Bol	.644 (.047)	.633 (.040)	.624 (.034)	.622 (.037)	.624 (.034)	.624 (.034)	.624 (.048)	.665 (.017)	.643 (.015)	.709 (.026)	.685 (.010)	.666 (.015)	.651 (.001)	.660 (.026)	.724 (.005)	.673 (.015)
 BoJ	.767 (.065)	.779 (.036)	.751 (.054)	.747 (.046)	.775 (.047)	.786 (.031)	.784 (.046)	.621 (.025)	.553 (.028)	.625 (.042)	.652 (.029)	.512 (.045)	.604 (.048)	.670 (.015)	.782 (.009)	.718 (.021)
 BoK	.801 (.049)	.837 (.034)	.853 (.018)	.779 (.108)	.828 (.029)	.771 (.105)	.844 (.048)	.658 (.026)	.561 (.054)	.650 (.030)	.721 (.031)	.582 (.042)	.637 (.027)	.686 (.020)	.801 (.008)	.760 (.019)
 BdeM	.651 (.017)	.620 (.053)	.673 (.094)	.682 (.035)	.656 (.062)	.702 (.054)	.625 (.048)	.666 (.055)	.541 (.027)	.685 (.023)	.703 (.035)	.454 (.007)	.575 (.040)	.629 (.007)	.750 (.012)	.639 (.051)
 BoT	.520 (.013)	.563 (.038)	.589 (.023)	.567 (.051)	.552 (.052)	.539 (.051)	.530 (.087)	.590 (.021)	.541 (.045)	.571 (.028)	.600 (.023)	.571 (.043)	.564 (.036)	.623 (.024)	.662 (.011)	.630 (.031)
 CBCT	.680 (.039)	.721 (.027)	.792 (.021)	.794 (.008)	.819 (.007)	.792 (.010)	.805 (.030)	.780 (.031)	.719 (.033)	.785 (.036)	.800 (.026)	.792 (.026)	.757 (.025)	.652 (.020)	.784 (.039)	.756 (.011)
 CBE	.842 (.015)	.830 (.038)	.854 (.053)	.858 (.024)	.858 (.011)	.822 (.016)	.863 (.046)	.829 (.031)	.815 (.014)	.862 (.020)	.852 (.024)	.839 (.015)	.838 (.026)	.806 (.017)	.904 (.027)	.874 (.021)
 CBR	.880 (.013)	.896 (.022)	.919 (.022)	.884 (.028)	.909 (.012)	.908 (.022)	.916 (.028)	.701 (.043)	.662 (.029)	.778 (.030)	.767 (.055)	.807 (.054)	.713 (.037)	.817 (.014)	.830 (.041)	.797 (.024)
 CBRT	.891 (.011)	.898 (.030)	.945 (.014)	.916 (.018)	.914 (.030)	.938 (.016)	.937 (.015)	.851 (.021)	.780 (.011)	.884 (.023)	.884 (.014)	.843 (.007)	.843 (.010)	.792 (.009)	.906 (.019)	.870 (.014)
 CBoC	.806 (.034)	.775 (.041)	.793 (.037)	.794 (.031)	.823 (.054)	.820 (.011)	.834 (.066)	.754 (.041)	.577 (.048)	.764 (.045)	.741 (.038)	.623 (.068)	.682 (.060)	.713 (.019)	.816 (.028)	.705 (.042)
 ECB	.838 (.027)	.878 (.052)	.888 (.041)	.877 (.036)	.888 (.025)	.873 (.015)	.876 (.011)	.735 (.019)	.645 (.006)	.722 (.027)	.762 (.023)	.669 (.039)	.735 (.010)	.701 (.036)	.841 (.007)	.787 (.013)
 FOMC	.773 (.025)	.777 (.008)	.818 (.025)	.803 (.011)	.826 (.023)	.824 (.008)	.837 (.021)	.818 (.019)	.763 (.045)	.806 (.016)	.779 (.034)	.718 (.048)	.818 (.032)	.671 (.033)	.839 (.022)	.825 (.011)
 MAS	.803 (.050)	.825 (.057)	.852 (.019)	.816 (.017)	.833 (.031)	.846 (.045)	.843 (.020)	.798 (.043)	.803 (.027)	.822 (.036)	.828 (.035)	.809 (.024)	.800 (.029)	.602 (.039)	.858 (.035)	.857 (.031)
 NBP	.803 (.020)	.803 (.029)	.855 (.009)	.849 (.026)	.814 (.006)	.840 (.017)	.844 (.021)	.760 (.013)	.610 (.021)	.725 (.007)	.745 (.023)	.623 (.026)	.689 (.033)	.607 (.023)	.813 (.016)	.774 (.019)
 PBoC	.929 (.005)	.943 (.012)	.941 (.014)	.928 (.006)	.933 (.006)	.937 (.012)	.936 (.004)	.641 (.046)	.749 (.017)	.851 (.013)	.801 (.023)	.869 (.022)	.844 (.023)	.927 (.021)	.955 (.010)	.942 (.012)
 RBA	.861 (.041)	.830 (.054)	.888 (.060)	.878 (.043)	.897 (.048)	.887 (.024)	.885 (.065)	.847 (.015)	.782 (.021)	.851 (.031)	.852 (.032)	.778 (.019)	.824 (.029)	.671 (.039)	.854 (.023)	.847 (.019)
 RBI	.754 (.028)	.780 (.075)	.812 (.040)	.813 (.080)	.829 (.007)	.813 (.099)	.836 (.068)	.662 (.018)	.584 (.017)	.700 (.024)	.686 (.021)	.616 (.020)	.599 (.042)	.750 (.030)	.755 (.023)	.702 (.029)
 SNB	.815 (.021)	.813 (.012)	.798 (.041)	.860 (.019)	.806 (.056)	.824 (.035)	.822 (.010)	.703 (.017)	.680 (.001)	.779 (.007)	.759 (.010)	.721 (.015)	.651 (.010)	.733 (.008)	.805 (.006)	.779 (.025)
Average	.785 (.097)	.797 (.093)	.820 (.093)	.808 (.093)	.819 (.092)	.811 (.019)	.820 (.100)	.737 (.077)	.677 (.088)	.764 (.079)	.764 (.067)	.711 (.109)	.713 (.087)	.702 (.093)	.816 (.066)	.776 (.078)

Table 17: F1-Scores across all three tasks (Stance Detection, Temporal Classification, and Uncertainty Estimation) are reported using the best-performing LLM on average, Llama-3-70B-Chat. This is evaluated under Few-Shot prompting and prompting with Annotation Guides. Standard deviations are shown in parentheses.

Bank	Few-Shot Prompting			Annotation-Guide Few-Shot Prompting		
	Stance	Temporal	Certain	Stance	Temporal	Certain
 BCB	.627 (.068)	.757 (.047)	.780 (.104)	.524 (.105)	.836 (.035)	.767 (.071)
 BCRP	.641 (.080)	.809 (.061)	.878 (.307)	.627 (.094)	.869 (.052)	.896 (.262)
 BNM	.577 (.129)	.865 (.037)	.745 (.104)	.588 (.080)	.875 (.039)	.762 (.065)
 BSP	.586 (.106)	.902 (.029)	.885 (.162)	.562 (.145)	.920 (.028)	.857 (.159)
 BanRep	.628 (.056)	.858 (.048)	.829 (.164)	.563 (.084)	.869 (.046)	.856 (.119)
 BoC	.661 (.195)	.820 (.048)	.624 (.092)	.718 (.075)	.869 (.031)	.769 (.055)
 BoE	.682 (.063)	.757 (.074)	.728 (.065)	.620 (.120)	.762 (.077)	.719 (.072)
 BoI	.585 (.159)	.882 (.033)	.729 (.143)	.427 (.135)	.844 (.032)	.695 (.205)
 BoJ	.583 (.104)	.718 (.023)	.784 (.167)	.616 (.078)	.781 (.022)	.808 (.101)
 BoK	.655 (.058)	.884 (.017)	.860 (.100)	.586 (.135)	.867 (.021)	.843 (.085)
 BdeM	.528 (.104)	.851 (.028)	.657 (.089)	.678 (.160)	.861 (.029)	.722 (.079)
 BoT	.573 (.188)	.723 (.067)	.671 (.083)	.596 (.213)	.711 (.050)	.684 (.070)
 CBCT	.611 (.147)	.840 (.024)	.665 (.118)	.542 (.087)	.833 (.016)	.745 (.056)
 CBE	.641 (.158)	.934 (.019)	.859 (.171)	.808 (.063)	.940 (.017)	.851 (.127)
 CBR	.731 (.142)	.846 (.044)	.905 (.132)	.784 (.127)	.826 (.038)	.909 (.123)
 CBRT	.647 (.085)	.847 (.049)	.884 (.148)	.645 (.110)	.934 (.031)	.901 (.132)
 CBoC	.663 (.160)	.788 (.041)	.848 (.140)	.698 (.075)	.792 (.048)	.824 (.109)
 ECB	.627 (.096)	.769 (.234)	.839 (.170)	.678 (.057)	.806 (.316)	.874 (.119)
 FOMC	.648 (.052)	.873 (.025)	.799 (.099)	.663 (.044)	.878 (.029)	.792 (.094)
 MAS	.663 (.128)	.893 (.019)	.783 (.115)	.583 (.148)	.940 (.010)	.814 (.076)
 NBP	.617 (.126)	.900 (.007)	.761 (.044)	.636 (.237)	.891 (.024)	.795 (.027)
 PBoC	.639 (.110)	.909 (.268)	.954 (.158)	.589 (.108)	.914 (.030)	.916 (.216)
 RBA	.572 (.174)	.933 (.033)	.815 (.119)	.636 (.192)	.921 (.032)	.862 (.085)
 RBI	.599 (.093)	.800 (.027)	.837 (.128)	.589 (.050)	.788 (.016)	.766 (.124)
 SNB	.648 (.095)	.824 (.023)	.782 (.126)	.613 (.090)	.853 (.021)	.797 (.098)
Average	.625 (.115)	.839 (.053)	.796 (.130)	.623 (.112)	.855 (.044)	.809 (.109)

Performance Analysis with Few-Shot and Annotation Guide Prompting. Table 17 shows the effect of Few-Shot and Annotation Guide prompting on Llama-3-70B-Chat. For Uncertainty Estimation, performance drops to 0.796 (± 0.130) with Few-Shot prompting and 0.809 (± 0.109) with Annotation Guide prompting. Temporal Classification shows a similar decline: 0.839 (± 0.053) with Few-Shot and 0.855 (± 0.044) with Annotation Guide. In contrast, Stance Detection shows a slight improvement, with scores increasing to 0.625 (± 0.115) for Few-Shot and 0.623 (± 0.112) for Annotation Guide. This translates to relative gains of 0.5% and 0.3% respectively. These results reinforce that Stance Detection remains the most challenging of the three tasks, where improved prompting offers minor but consistent gains. However, even with enhanced prompting, LLMs do not surpass PLM performance, suggesting that the performance ceiling for LLMs remains lower than that of fine-tuned PLMs on the Stance Detection task.

1071 C.6 Whole Model Performance Gain

1072 In our experiments, we discover that there is a significant improvement in model performance across
 1073 all three labels when the training corpus of the is expanded from just the specific bank (Bank Specific
 1074 Setup) to having all banks' data (General Setup). In order to comprehend why this happens, we
 1075 evaluate the models' increased performance using qualitative and statistical reasoning.

1076 C.6.1 Qualitative Reasoning

1077 To understand the improved performance of the model trained on the whole corpus of the annotated
 1078 data as compared to the specific banks' models, we qualitatively evaluate the Stance Detection

1079 labels given by each model and compare them to the ground truth (Table 18). For each bank, we
1080 compare its best Bank Specific Setup model against our best General Setup model. This comparison
1081 demonstrates the general model benefits from patterns between central bank communications, which
1082 suggests that the model is creating meaningful parallels across central banks. These findings highlight
1083 the potential for inter-bank learning, where information about one central bank’s communications
1084 improves the model’s ability to interpret another bank’s communications.

1085 We then further generate embeddings for each sentence within our training corpus and find the
1086 sentence whose embeddings are the closest to our selected sentence as well as its label and origin
1087 bank as seen within Table 18. We find that there is a significant amount of learning from other banks.
1088 When the model is trained on a larger corpus, it is able to learn from the annotations and labels of the
1089 other banks.

1090 C.6.2 Statistical Reasoning

1091 To further understand this improvement in model performance from the specific model to the
1092 generalized model trained on all the annotated sentences for all central banks, we hypothesize that
1093 there are semantic similarities across central bank meeting level documents that enable the general
1094 setup model to generalize effectively. To test our hypothesis, we perform TF-IDF for every central
1095 bank at a document level and then using cosine similarity to find the level of similarity between
1096 banks’ documents. After normalizing these scores and setting them up as the independent variable,
1097 we compare them using a linear regression with the change in the weighted F1-Score of the best
1098 performing model (RoBERTa-large) on average for both specific and general cases. This then leads
1099 to the following formula:

$$y = 0.0168x_1 + 0.0511$$

1100 This linear regression has a p-value of 0.016, implying statistical significance under the threshold of
1101 2%. These results indicate that semantic overlap between banks’ document level data contributes
1102 to a general improvement in model performance, supporting the hypothesis that central bank com-
1103 munications contain consistencies that allow models to learn from other banks’ data to improve
1104 performance.

Table 18: Comparison of the stance labels generated by the whole and specific model (using seed = 5768) for each central bank. The closest bank, sentence and label columns represent the results from generating and comparing embeddings from the training corpus. * indicates that the general model's label was the ground truth.

Bank	Sentence	Specific model	General Model*	Closest Bank	Closest Sentence	Closest Label
 CBoC	although establishing such a conclusion required reviewing multiple dimensions, some of them could be mentioned.	Neutral	Irrelevant	 CBCT	related discussions are summarized as follows	Irrelevant
 BoJ	meanwhile, business sentiment had generally stayed at a favorable level, although some cautiousness had been observed.	Dovish	Neutral	 CBE	moreover, most leading indicators remained in positive territory in 2022 q2.	Hawkish
 CBCT	related discussions are summarized as follows	Neutral	Irrelevant	 CBoC	although establishing such a conclusion required reviewing multiple dimensions, some of them could be mentioned.	Irrelevant
 BdeM	one member emphasized that its seasonally adjusted and annualized monthly rate was 3.53% in July 2024.	Hawkish	Neutral	 NBP	other council members assessed that the recent decline in service price growth had probably been temporary	Neutral
 ECB	US Commodity Futures Trading Commission numbers suggested that the market remained positioned for an appreciation of the euro.	Hawkish	Dovish	 FOMC	the nominal deficit on U.S. trade in goods and services continued to widen in April.	Hawkish
 BoE	there were also arguments in favour of a programme of purchases towards the upper end of the range	Irrelevant	Dovish	 NBP	other council members assessed that the recent decline in service price growth had probably been temporary	Neutral
 CBE	moreover, most leading indicators remained in positive territory in 2022 q2.	Neutral	Hawkish	 BoJ	meanwhile, business sentiment had generally stayed at a favorable level, although some cautiousness had been observed.	Neutral
 BNM	in line with earlier assessments, headline inflation is likely to have peaked in 3q 2022 and is expected to moderate thereafter, albeit remaining elevated	Dovish	Hawkish	 CBE	moreover, most leading indicators remained in positive territory in 2022 q2.	Hawkish
 MAS	consequently, unit labour costs will increase at a significantly slower pace compared with the preceding two years.	Neutral	Dovish	 BoK	increase in the prices of petroleum products and the base effect from the decline in the prices of public services last year.	Hawkish
 BanRep	in the same period, other commodities exported and imported by Colombia descended.	Neutral	Dovish	 SNB	what the economy needs to recover, however, is for exports to pick up	Dovish
 BoT	third, the continued economic recovery of trading partners would benefit Thai exports.	Neutral	Dovish	 SNB	what the economy needs to recover, however, is for exports to pick up	Dovish
 FOMC	the nominal deficit on U.S. trade in goods and services continued to widen in April.	Neutral	Hawkish	 BoK	increase in the prices of petroleum products and the base effect from the decline in the prices of public services last year.	Hawkish
 PBoC	the PBC will implement monetary policy in a flexible and targeted manner, and strengthen the coordination with fiscal, industrial, and regulatory policies.	Neutral	Dovish	 BoI	in the last twelve months the CPI has increased by 2.5 percent, within the target inflation range.	Neutral
 RBI	the latest assessment by the World Trade Organisation (WTO) for Q4 indicates a loss of momentum in global trade due to declining export orders.	Hawkish	Dovish	 FOMC	the nominal deficit on U.S. trade in goods and services continued to widen in April.	Hawkish
 BoK	increase in the prices of petroleum products and the base effect from the decline in the prices of public services last year.	Dovish	Hawkish	 BoI	in the last twelve months the CPI has increased by 2.5 percent, within the target inflation range.	Neutral
 SNB	what the economy needs to recover, however, is for exports to pick up	Hawkish	Dovish	 BoT	third, the continued economic recovery of trading partners would benefit Thai exports.	Dovish
 CBRT	housing loan rates, which have remained flat in the same period, stood at 41.4%.	Irrelevant	Neutral	 CBR	unsecured consumer lending is supported by higher households' incomes.	Dovish
 BCRP	these measures have allowed immediate release of funds in domestic currency amounting to a total of 2 billion	Neutral	Dovish	 BCB	the result was creation of 590 thousand jobs in the first half of the year	Hawkish
 BSP	the next meeting of the Monetary Board to discuss monetary policy is scheduled on 12 January 2006.	Neutral	Irrelevant	 PBoC	the PBC will implement monetary policy in a flexible and targeted manner, and strengthen coordination with fiscal, industrial, and regulatory policies.	Dovish
 BCB	the result was creation of 590 thousand jobs in the first half of the year	Neutral	Hawkish	 BCRP	these measures have allowed immediate release of funds in domestic currency amounting to a total of 2 billion	Dovish
 RBA	however, central banks had also emphasised that policy rates were unlikely to decline over coming months, in contrast to market-implied expectations.	Dovish	Hawkish	 NBP	other council members assessed that the recent decline in service price growth had probably been temporary	Neutral
 NBP	other council members assessed that the recent decline in service price growth had probably been temporary	Hawkish	Neutral	 BoK	increase in the prices of petroleum products and the base effect from the decline in the prices of public services last year.	Hawkish
 CBR	unsecured consumer lending is supported by higher households' incomes.	Neutral	Dovish	 CBRT	housing loan rates, which have remained flat in the same period, stood at 41.4%.	Neutral
 BoI	in the last twelve months the CPI has increased by 2.5 percent, within the target inflation range.	Hawkish	Neutral	 BoK	increase in the prices of petroleum products and the base effect from the decline in the prices of public services last year.	Hawkish
 BoC	GDP growth is forecast to increase in the second half of 2024 and through 2025	Neutral	Hawkish	 SNB	what the economy needs to recover, however, is for exports to pick up	Dovish

1105 C.7 Experiment: Training Set Size

1106 To assess how the stance detection label model trained on the full corpus behaves across varying
 1107 amounts of training data, we ran experiments with sample sizes ranging from 40 to 700. As shown in
 1108 Figure 11 below, the F1-Score begins to plateau at 0.74, indicating that changing the sample size does
 1109 not impact performance after 600 samples. These experiments for the Stance Detection label
 1110 were run on the RoBERTa-Large model l using seeds = [5768, 78516, 944601], sample
 1111 sizes = 32 and learning rates = 1e-6

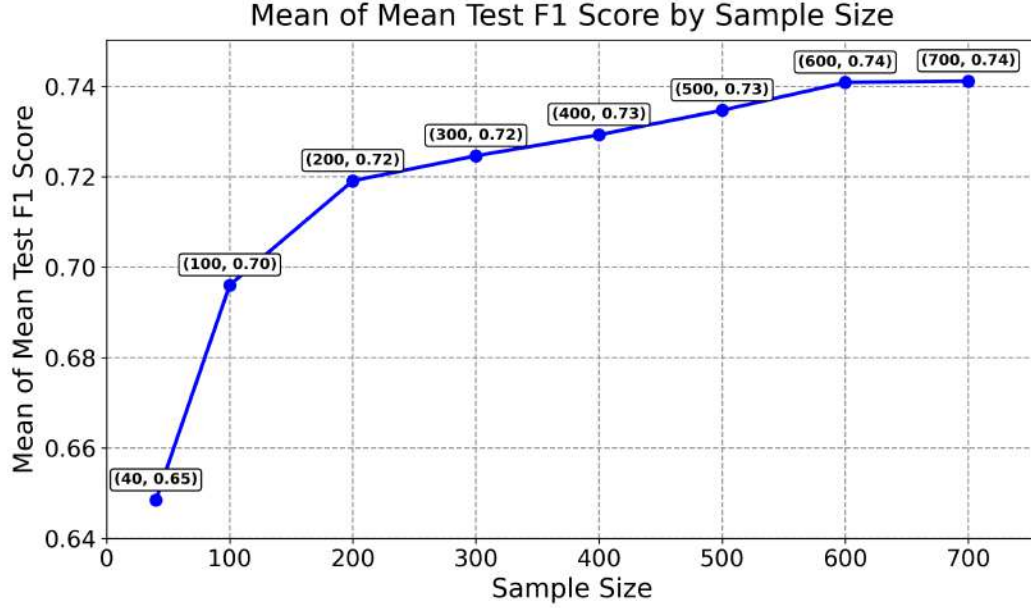


Figure 11: Impact of varying training sample sizes on average test F1-Score of best performing model using the entire annotated dataset. Performance rapidly improves until 600 samples, after which performance starts to plateau.

1112 C.8 Hyperparameters

1113 C.8.1 WCB Models

1114 Table 19 displays the best performing hyperparameters for each task given the General Setup. We
 1115 obtain these hyperparameters through grid search across the following parameters: seeds = [5768,
 78516, 944601]², batch sizes = [32, 16], learning rates = [1e-5, 1e-6].

Table 19: Best model hyperparameter for each task for the model trained using the General Setup.

Task	Model Name	Learning Rate	Batch Size	Best Seed
Certain Label	roberta-large	1e-05	16	5768
Stance Label	roberta-large	1e-06	32	944601
Time Label	roberta-base	1e-06	32	944601

1116


























1117 C.8.2 Specific Bank Models

1118 For each bank and task under the specific setup, the hyperparameters that have the best performance
 1119 for the given top-ranked PLM are shown in the tables below. Table 20 show the best hyperpa-
 1120 rameters for the best performing model for Uncertainty Estimation task while Table 21 and

²These specific seeds were chosen based on the work of Shah et al. [2022] and Shah et al. [2023].

1121 Table 22 show the best hyperparameters for the best performing model for Stance Detection and
 1122 Temporal Classification, respectively. We obtain these hyperparameters through grid search
 1123 across the following parameters: seeds = [5768, 78516, 944601]³, batch sizes = [32,
 16], learning rates = [1e-5, 1e-6].

Table 20: Best performing model and corresponding hyperparameter-learning rate and batch size for Uncertainty Estimation.

Bank	Base Model	Learning Rate	Batch Size	Best Seed	Link
 FOMC	roberta-large	1e-06	16	78516	HF Link
 PBoC	ModernBERT-large	1e-05	16	78516	HF Link
 BoJ	roberta-base	1e-05	16	944601	HF Link
 BoE	bert-base	1e-06	16	944601	HF Link
 SNB	bert-large	1e-05	16	5768	HF Link
 BCB	bert-base	1e-05	32	78516	HF Link
 RBI	roberta-large	1e-05	16	944601	HF Link
 ECB	finbert-pretrain	1e-06	16	5768	HF Link
 CBR	bert-base	1e-05	32	5768	HF Link
 CBCT	finbert-pretrain	1e-06	16	78516	HF Link
 MAS	bert-base	1e-06	16	944601	HF Link
 BoK	bert-base	1e-06	16	5768	HF Link
 RBA	finbert-pretrain	1e-05	16	944601	HF Link
 BoI	ModernBERT-base	1e-05	32	944601	HF Link
 BoC	bert-large	1e-05	16	944601	HF Link
 BdeM	roberta-base	1e-06	16	5768	HF Link
 NBP	bert-base	1e-05	32	78516	HF Link
 CBRT	bert-base	1e-05	32	78516	HF Link
 BoT	bert-base	1e-05	32	78516	HF Link
 CBE	roberta-large	1e-06	16	5768	HF Link
 BNM	roberta-large	1e-05	32	78516	HF Link
 BSP	ModernBERT-base	1e-05	16	5768	HF Link
 CBoC	roberta-large	1e-05	16	78516	HF Link
 BCRP	finbert-pretrain	1e-05	32	78516	HF Link
 BanRep	roberta-large	1e-06	32	78516	HF Link

1124

³These specific seeds were chosen based on the work of Shah et al. [2022] and Shah et al. [2023].

Table 21: Best performing model and corresponding hyperparameter-learning rate and batch size for Stance Detection.



















































Bank	Base Model	Learning Rate	Batch Size	Best Seed	Link
 FOMC	roberta-large	1e-06	32	944601	HF Link
 PBoC	roberta-large	1e-05	16	78516	HF Link
 BoJ	roberta-base	1e-05	32	78516	HF Link
 BoE	roberta-large	1e-06	16	5768	HF Link
 SNB	roberta-large	1e-06	16	78516	HF Link
 BCB	roberta-large	1e-05	16	944601	HF Link
 RBI	roberta-large	1e-06	32	944601	HF Link
 ECB	roberta-large	1e-06	32	5768	HF Link
 CBR	roberta-large	1e-05	16	5768	HF Link
 CBCT	roberta-large	1e-06	16	5768	HF Link
 MAS	roberta-base	1e-06	16	78516	HF Link
 BoK	roberta-large	1e-06	32	5768	HF Link
 RBA	roberta-base	1e-05	16	78516	HF Link
 BoI	roberta-large	1e-06	16	78516	HF Link
 BoC	roberta-large	1e-06	16	5768	HF Link
 BdeM	roberta-large	1e-06	16	5768	HF Link
 NBP	roberta-large	1e-05	16	944601	HF Link
 CBRT	roberta-large	1e-06	32	944601	HF Link
 BoT	roberta-large	1e-05	32	78516	HF Link
 CBE	roberta-large	1e-05	16	78516	HF Link
 BNM	ModernBERT-large	1e-05	32	5768	HF Link
 BSP	roberta-large	1e-06	16	944601	HF Link
 CBoC	roberta-large	1e-05	16	944601	HF Link
 BCRP	roberta-large	1e-06	16	5768	HF Link
 BanRep	roberta-large	1e-05	16	78516	HF Link

Table 22: Best performing model and corresponding hyperparameter-learning rate and batch size for Temporal Classification.

Bank	Base Model	Learning Rate	Batch Size	Best Seed	Link
 FOMC	roberta-large	1e-05	32	5768	HF Link
 PBoC	ModernBERT-base	1e-05	16	944601	HF Link
 BoJ	roberta-base	1e-06	32	944601	HF Link
 BoE	roberta-base	1e-05	32	944601	HF Link
 SNB	roberta-large	1e-06	16	5768	HF Link
 BCB	roberta-large	1e-05	16	944601	HF Link
 RBI	roberta-base	1e-06	16	944601	HF Link
 ECB	roberta-base	1e-06	16	78516	HF Link
 CBR	bert-base	1e-06	16	5768	HF Link
 CBCT	bert-large	1e-05	16	944601	HF Link
 MAS	finbert-pretrain	1e-05	32	5768	HF Link
 BoK	roberta-large	1e-06	32	5768	HF Link
 RBA	roberta-large	1e-05	16	78516	HF Link
 BoI	finbert-pretrain	1e-05	32	944601	HF Link
 BoC	roberta-base	1e-06	32	944601	HF Link
 BdeM	roberta-base	1e-06	32	944601	HF Link
 NBP	roberta-large	1e-05	16	78516	HF Link
 CBRT	roberta-base	1e-05	16	5768	HF Link
 BoT	roberta-large	1e-06	32	944601	HF Link
 CBE	ModernBERT-large	1e-05	16	78516	HF Link
 BNM	finbert-pretrain	1e-06	16	78516	HF Link
 BSP	roberta-large	1e-06	16	944601	HF Link
 CBoC	roberta-base	1e-06	32	944601	HF Link
 BCRP	roberta-large	1e-05	32	5768	HF Link
 BanRep	roberta-base	1e-06	16	5768	HF Link

1125 D Extended Experiments and Analysis

1126 In our entire study, we run 15,075 experiments. These include benchmarking across both setups for 25
1127 central banks and 16 models. The PLMs were benchmarked using gridsearch across `seeds = [5768,`
1128 `78516, 944601]`, `batch sizes = [32, 16]`, `learning rates = [1e-5, 1e-6]`. The
1129 LLMs were tested under a single set of hyperparameters but using the same 3 seeds. We also
1130 run few shot prompting and annotation guide prompting for the best performing LLM across all
1131 banks and the same 3 seeds. In addition to this, we conduct several experiments as described below.

1132 D.1 People’s Bank of China LLM Error Analysis

1133 In order to better understand how well our models classify statements across the three tasks (`Stance`
1134 `Detection`, `Temporal Classification`, and `Uncertainty Estimation`), we conduct an er-
1135 ror analysis of `Llama-3-70b-chat-hf`, the best-performing LLM on the PBoC dataset. We perform
1136 error analysis on all the sentences misclassified for `Temporal Classification` and `Uncertainty`
1137 `Estimation`. For `Stance Detection`, we randomly sample 50 statements from the 150 mislabeled
1138 sentences with `random seed=1`. A sample of mislabeled sentences and the corresponding reason
1139 can be viewed in Table D.1.3

1140 D.1.1 Stance Detection

1141 For `Stance Detection`, we identify three common mistakes: (1) **Leaning Dovish** (32/50): The
1142 model fails to recognize when a statement is slightly Dovish and incorrectly labels such statements
1143 as `Neutral`. (2) **Misidentifies monetary policy** (14/50): The LLM misidentifies statements that
1144 advocate for implementation of monetary policy as `Neutral` or `Irrelevant` and statements that do
1145 not advocate for implementation of monetary policy as `Hawkish` or `Dovish`. (3) **Misunderstanding**
1146 (4/50): The model misunderstands the statement entirely.

1147 D.1.2 Temporal Classification

1148 For `Temporal Classification`, we identify four common errors: (1) **Changing trends** (20/34):
1149 The LLM mislabels statements about changing trends in the economy as `Not Forward Looking`
1150 instead of `Forward Looking`. (2) **Current event** (5/34): The LLM mistakenly labels sentences
1151 about current events as `Not Forward Looking`. (3) **Past tense** (3/34): The LLM misidentifies
1152 statements written in past tense as `Forward Looking`. (4) **Misunderstanding** (11/34): The model
1153 misunderstands the statement entirely.

1154 D.1.3 Uncertainty Estimation

1155 For `Uncertainty Estimation`, we identify three common errors: (1) **Known economic trends**
1156 (7/20): The LLM misidentifies sentences that state the economic trends with certainty as `Uncertain`.
1157 (2) **Uncertain actions** (5/20): The LLM incorrectly labels sentences with no clear monetary policy
1158 as `Certain`. (3) **Misunderstanding** (8/20): The model misunderstands the statement entirely.

1159 D.2 Generating Meeting Minutes

1160 As discussed in Section 1, central bank communications shape expectations about economic stability.
1161 To this end, having an early sense of their stance can benefit the broader public. Using our dataset,
1162 we explore whether a long-context LLM can generate a target central bank’s next monetary policy
1163 document (minutes or their equivalent). The model receives only (i) the target bank’s own last
1164 document issued before that meeting (its most recent pre-cutoff release) and (ii) documents from
1165 the other 24 banks, but only those released prior to the target document’s release date. We then test
1166 whether the policy stance of the generated document matches that of the actual target document.
1167 Alignment is measured with the hawkishness metric defined in Section 6, and our fine-tuned `Stance`
1168 `Detection` model provides the evaluation benchmark.

1169 D.2.1 Methodology

1170 We generate candidate meeting documents (minutes or their equivalent) with GPT-4.1, chosen for
1171 its one million token context window and top performance among the long-context models in our

Table 23: Sample of sentences from error analysis. *Correct label* is the actual label for that sentence. *LLM Label* is the label that the Llama-3 model predicted. *Type of Error* is the error that the LLM made.

Sentence	Correct Label	LLM Label	Type of Error
it was stressed that the sound monetary policy should be continued and the Forward Looking, scientific and effective approach to macro financial management be further strengthened in the coming period.	Dovish	Neutral	Leaning dovish
however, economic stabilization and recovery is not firmly established.	Neutral	Dovish	Misidentifies monetary policy
the principle of making the rmb exchange rate reform a self-initiated, controllable and gradual process should be followed to keep the rmb exchange rate basically stable at an adaptive and equilibrium level.	Forward Looking	Not Forward Looking	Changing trends
the overall performance of the financial sector has remained sound.	Forward Looking	Not Forward Looking	Current event
however, the chinese economy still faced a complex situation, with both favorable and unfavorable factors in sight.	Not Forward Looking	Forward Looking	Past tense
the solution of the problem needs cooperation among the enterprises, the commercial banks as well as authorities in different sectors.	Certain	Uncertain	Known economic trends
while financial support to economic growth must be guaranteed, no negligence should be tolerated in preventing inflation and financial risks.	Uncertain	Certain	Uncertain actions

benchmark. For each target central bank, we provide documents from the other 24 banks and query the model to produce the target bank’s first post-cutoff document. We then compute hawkishness scores for the generated text (\hat{z}_t) and the actual document’s (z_t) using the best-performing fine-tuned **Stance Detection** model. A linear regression of \hat{z}_t on z_t across all banks quantifies how well the generated stance tracks the actual stance in the target document. The full procedure is:

1. Let $\mathcal{B} = \{b_1, b_2, \dots, b_{25}\}$ denote the set of 25 central banks in our dataset.
2. For each target central bank $b_t \in \mathcal{B}$, let Y_t be its first released document after the fixed cutoff date for GPT 4.1 (May 31, 2024). This is treated as the **actual output**.
3. Let Y'_t be the most recent document from b_t released *before* Y_t . This is used as the **example output**.
4. For the **example input set** X' , we collect all documents from the remaining 24 banks $\mathcal{B} \setminus \{b_t\}$ that were released before the timestamp of Y'_t . This yields a globally available context prior to Y'_t .
5. For the **actual input set** X , we collect all documents from $\mathcal{B} \setminus \{b_t\}$ released before Y_t . It is important to note that there may be partial overlap between X and X' depending on each central bank’s document frequency.
6. The prompt is constructed with (X', Y'_t) as an example and (X) as actual input. The LLM is then queried to generate \hat{Y}_t , the **predicted meeting document** for b_t . The prompt structure is illustrated in Figure 12.
7. We compute the **hawkishness measure** for each document Y_t and the generated document \hat{Y}_t , denoted z_t and \hat{z}_t respectively. This is done using the hawkishness measure described in Section 6.
8. Finally, we fit a linear regression model $\hat{z}_t = \alpha + \beta z_t + \epsilon$ across all central banks to test how closely the stance predicted from the generated minutes follows the stance of the corresponding actual document. We use root mean square error (RMSE) as the primary metric and report full regression statistics.

D.2.2 Hawkishness Measure

To ensure a fair evaluation, we use our best-performing fine-tuned **Stance Detection** model (RoBERTa-Large) to label sentences in both the generated and actual documents as Hawkish, Dovish, Neutral, or Irrelevant (Section 6). We then compute the document-level hawkishness



Figure 12: Prompt used to generate meeting minutes using GPT 4.1. Orange examples show how placeholders are populated at inference time. The example sentences shown are illustrative and do not reflect actual central bank communications.

1202 score using the previously defined formula:

$$z(D) = \frac{\# \text{Hawkish Sentences} - \# \text{Dovish Sentences}}{\# \text{Total Sentences} - \# \text{Irrelevant Sentences}},$$

1203 where counts are taken over all sentences in document D . We write z_t for the true post-cutoff minutes

1204 Y_t and \hat{z}_t for the generated minutes \hat{Y}_t .

To assess how well the LLM preserves policy stance, we fit the linear regression

$$\hat{z}_t = \alpha + \beta z_t + \epsilon$$

across all 25 banks. We obtain $\beta = 0.6306$ ($p \leq 10^{-4}$), indicating a highly significant relationship and confirming that higher true hawkishness reliably predicts higher generated hawkishness.

D.3 Human Evaluation

We aim to evaluate the ability of our models to understand central communications with significant real-world consequences. We compare the best-performing LLM for each task (Stance Detection, Temporal Classification and Uncertainty Estimation) based on weighted F1-Score on the European Central Bank (ECB) dataset to that of a non-expert annotator with no annotation guide. We specifically chose ECB because its policy moves anchor rates for 20 member states and greatly impact global fixed-income and foreign exchange markets. The annotator is a PhD-candidate with formal work experience within the financial domain and no prior knowledge of European Union (EU) monetary policy. The best performing LLM for Stance Detection and Temporal Classification is Llama-3-70b-chat and gemini-2.0-flash for Certainty Estimation. We compare the LLM and human annotations to the ground truth. The results are summarized in Tables 14, 13, and 15.

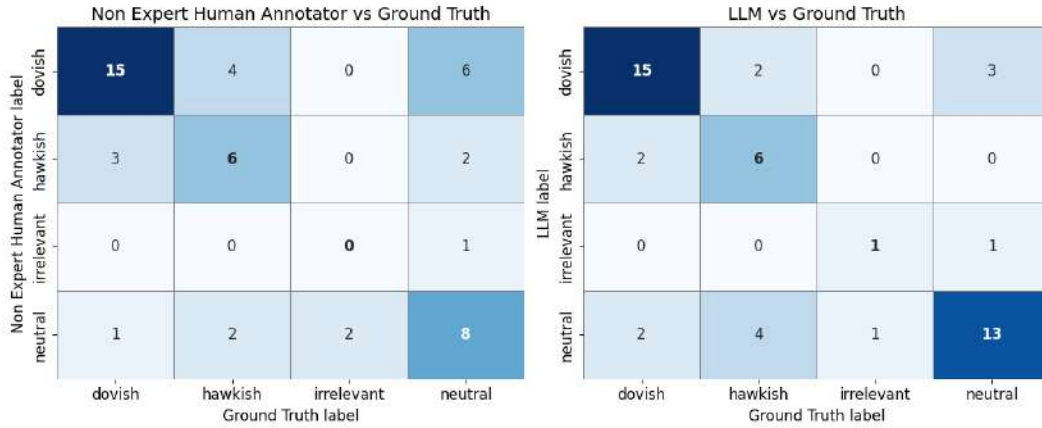


Figure 13: Side-by-side confusion matrices comparing a non-expert human annotator (left) and a large language model (right) against ground-truth stance labels: Dovish, Hawkish, Irrelevant, and Neutral. Each cell indicates the number of sentences for each ground-truth (row) and predicted (column) label pair (e.g. four Hawkish sentences were labeled as Neutral by the LLM).

Across all three tasks, the LLM consistently makes fewer mistakes than the non-expert human annotator. For Stance Detection, the non-expert correctly classifies 29 sentences whereas the LLM classifies 35 sentences correctly, with most human errors being confusing Hawkish and Neutral sentences. For Uncertainty Estimation, the human mislabeled 12 Uncertain sentences as Certain while the LLM made 11 such errors. For Temporal Classification, the human mislabeled 10 Forward Looking sentences as Not Forward Looking compared to only four such mistakes by the LLM. These results highlight the inherent difficulty of discerning subtle cues, especially for Neutral or nuanced Forward Looking sentences in the case of the ECB dataset; and our models' emerging strength at maintaining consistency across complex, domain-specific labeling tasks.

D.4 Transfer Learning

D.4.1 Beyond 25 Banks: Czech National Bank

For further test the ability of our best performing models (RoBERTa-Large for Stance Detection and Uncertainty Estimation, RoBERTa-Base for Temporal Classification) to generalize to different economies, we test our models on sanitized data from the Czech National Bank (CNB).

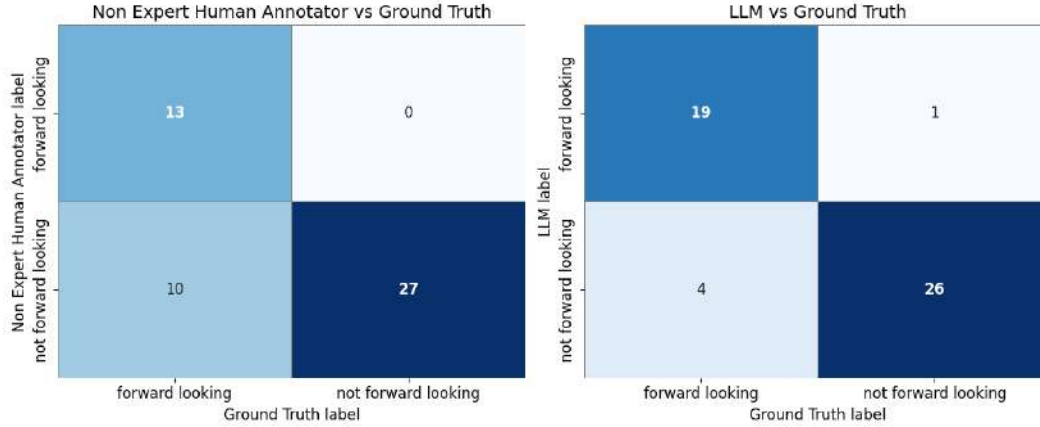


Figure 14: Side-by-side confusion matrices comparing a non-expert human annotator (left) and a LLM (right) against ground-truth temporal labels: Forward Looking, Not Forward Looking. Each cell indicates the number of sentences for each ground-truth (row) and predicted (column) label pair (e.g., 10 Forward Looking statements were labeled as Not Forward Looking by the LLM).

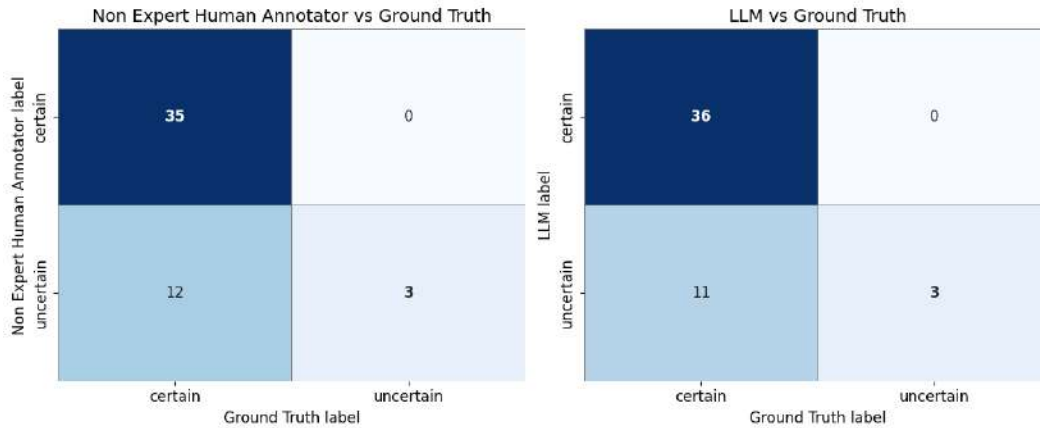


Figure 15: Side-by-side confusion matrices comparing a non-expert human annotator (left) and a LLM (right) against ground-truth certainty labels: Uncertain, Certain. Each cell indicates the number of sentences for each ground-truth (row) and predicted (column) label pair (e.g., 12 Certain statements were labeled as Uncertain looking by the LLM).

1235 As with all the other central bank data, we scrape, clean, and tokenize the CNB data from 1998 to
 1236 2024. We uniformly sample 500 sentences and use the corresponding annotation guide as seen in
 1237 Appendix H.26 to label these sentences. Then, we test our fine-tuned models on the labeled CNB
 1238 data and obtain the following weighted F1-Scores and their standard deviations:

- 1239 • **Stance Label:** 0.800(0.018)
- 1240 • **Uncertainty Estimation Task:** 0.702(0.010)
- 1241 • **Time Label:** 0.726(0.030)

1242 This demonstrates that our models perform well on data from central banks not present within our
 1243 corpus of 25 banks and highlight the generalizability of our models.

1244 D.4.2 Beyond Finance: The Future of Conservation Climate, Forestry, and Natural Resources

1245 Since Temporal Classification and Uncertainty Estimation tasks extend beyond the field
 1246 of economics, we test the applicability of our best performing models for each on the United States
 1247 Senate Committee on Agriculture, Nutrition, and Forestry hearing transcripts [USCongress, 2023].

1248 From USCongress [2023], we sample 114 sentences using selenium and sanitize them in the same
 1249 manner we did for the central bank communications. Using these sentences, we label each sentence
 1250 as Forward Looking Or Not Forward Looking and Certain or Uncertain using Table 24 and
 1251 25. This was done by an expert annotator with the help of an annotation guide. We test our fine-tuned
 1252 models on the labeled dataset. The results are summarized below.

ESG Annotation Guide

Table 24: Annotation guide for classifying ESG-related sentences as Forward Looking or Not Forward Looking.

Label	Description	Example
Forward Looking	When the sentence discusses expectations, projections, or anticipations about future environmental conditions, policy actions, or resource utilization.	“Moving forward, I hope that you will commit to increasing your support of the valuable locally led initiatives like those of our conservation partners and those in the Colorado Department of Agriculture that you will hear more about from our commissioner of ag.”
Not Forward Looking	When the sentence reflects on recent or past environmental decisions, trends, or events to describe what has already occurred.	“We do not have the wonderful forest that Colorado has for a carbon sink, but the soil carbon sinks that we have in our States are just as important.”

1253

Table 25: Annotation guide for classifying sentences from the Agriculture, Nutrition, and Forestry Committees as Certain or Uncertain

Label	Description	Example
Certain	When an expectation, trend, action, or outcome is definitively stated without ambiguity. Key words include “is,” “are,” “decreased,” and “will continue.”	“Last but not least, Mr. Alexander Funk is the director of Water Resources and senior counsel for the Theodore Roosevelt Conservation Partnership.”
Uncertain	When it expresses concerns, potential risks, or need to alter monetary policy if a specific event comes to pass. Key words include “can,” “if,” “could,” and “uncertain.”	“That said, the conservation title can play a great role in addressing drought conditions, and there are several opportunities to help Western farmers and ranchers.”

1254 **Results** As a result of the testing on the Congress Committee hearings data from the conservation,
 1255 climate, forestry, and natural resources of United States Senate Committee on Agriculture, Nutrition,
 1256 and Forestry, we find the following weighted F1-Scores:

- 1257 • **Time Label:** 0.879
- 1258
- 1259 • **Certain Label:** 0.683
- 1260

1261 Furthermore, according to Tables 26 and 27, the models correctly labeled 100/144(88%) sentences
 1262 for Temporal Classification and 87/114(76%) sentences for Uncertainty Estimation.

1263 We see that these scores are significant, demonstrating the applicability of our general models
 1264 fined-tuned for Temporal Classification and Uncertainty Estimation to other domains.

Table 26: Confusion matrix comparing predicted and ground truth Temporal Classification of ESG-related sentences.

Temporal Classification		
	Predicted Not Forward	Predicted Forward
True Not Forward	35	4
True Forward	10	65

Table 27: Confusion matrix comparing predicted and ground truth Uncertainty Estimation of sentences from Agriculture, Nutrition, and Forestry Committees.

Uncertainty Estimation		
	Predicted Certain	Predicted Uncertain
True Certain	84	0
True Uncertain	27	3

1265 E Extended Economic Analysis

1266 E.1 Data Collection and Aggregation

1267 We collect year-over-year (YoY), percent change data for non-seasonally adjusted Consumer Price
 1268 Index (CPI) from the Bloomberg Terminal at the Georgia Institute of Technology. We use the non-
 1269 seasonally adjusted indices because it offers a more complete picture of the macroeconomic trends at
 1270 that time. All central banks use CPI, with the exception of the European Central Bank and the Bank
 1271 of England. They use Harmonized CPI which is CPI excluding the rent-equivalent of owner-owned
 1272 dwellings.

The table below lists the exact CPI index used for each central bank.

Table 28: Summary of metadata for Consumer Price Index (CPI) data corresponding to the 25 central banks. *Name of Index* refers to the official title of the index. *Index* denotes the Bloomberg ticker. *Source* indicates the government agency responsible for data collection. *Date* and *Time Collected* specify when the data was retrieved from the Bloomberg Terminal.

Bank	Name of Index	Index	Source	Date lected	Col- lected	Time lected	Col- lected
 FOMC	US CPI Urban Consumers YoY NSA	CPI YOY	Bureau of Labor Statistics	Apr 22		15:07	
 PBoC	China CPI YoY	CNCPIYOY	National Bureau of Statistics of China	Apr 22		15:07	
 BoJ	Japan CPI Nationwide YoY	JNCPIYOY	Ministry of Internal Affairs and Communications	Apr 22		15:07	
 BoE	UK CPI EU Harmonized YoY NSA	UKRPCJYR	UK Office for National Statistics	Apr 22		15:07	
 SNB	Switzerland CPI All Items YoY	SZCPIYOY	Federal Statistics Office of Switzerland	Apr 22		15:07	
 BCB	Brazil CPI IPCA YoY	BZPIPCY	The Organization for Economic Co-operation and Development	Apr 22		15:07	
 RBI	India CPI Combined YoY	INFUTOTY	Central Statistics Office India	Apr 22		15:20	
 ECB	Eurostat EU HICP All Items YoY NSA	ECCPEUY	Eurostat	Apr 22		15:20	
 CBR	Russia CPI YoY	RUCPIYOY	Federal Service of State Statistics	Apr 22		15:20	
 CBCTW	Taiwan CPI YoY NSA	TWCPIYOY	Taiwan Directorate General of Budget Accounting and Statistics	Apr 22		15:20	
 MAS	Singapore CPI All Items YoY	SICPIYOY	Singapore Ministry of Trade and Industry	Apr 22		15:20	
 BoK	South Korea CPI YoY	KOCPIYOY	Bank of Korea	Apr 22		15:20	
 RBA	Australia CPI All Items YoY	AUCPIYOY	Australian Bureau of Statistics	Apr 22		15:20	
 BoI	Israel CPI YoY NSA	ISCPIYYN	Israel Central Bureau of Statistics	Apr 22		15:20	
 BoC	STCA Canada CPI YoY NSA 2002=100	CACPIYOY	Statistics Canada	Apr 22		15:20	
 BdeM	Mexico CPI Yearly Percent Change Biweekly	MXBWYOY	National Institute of Statistics and Geography	Apr 22		15:20	
 NBP	Poland CPI All Items YoY	POCPIYOY	Polish Statistics Office	Apr 22		15:20	
 CBRT	Turkey CPI YOY %	TUCPIY	Turkish Statistical Institute	Apr 22		15:20	
 BoT	Thailand CPI All Items YoY	THCPIYOY	Office of the National Economic and Social Development Council	Apr 22		15:20	
 BNM	Malaysia CPI YoY 2010=100	MACPIYOY	Department of Statistics Malaysia	Apr 22		15:20	
 BSP	Philippine CPI All items YoY% 2018=100	PHC2II	Philippine Statistics Authority	Apr 27		13:28	
 CBE	Egypt Monthly Urban CPI YoY%	EGCPYOY	Central Agency for Public Mobilization and Statistics	Apr 22		15:20	
 CBoC	Chile CPI YoY Chained	CLINNSYO	Instituto Nacional de Estadística – Chile	Apr 22		15:20	
 BCRP	Lima CPI YoY	PRCPYOY	Instituto Nacional De Estadística E Informatica De Peru	Apr 22		15:20	
 BanRep	Colombia CPI YoY	COCPIYOY	Departamento Administrativo Nacional de Estadística	Apr 22		15:20	

1273

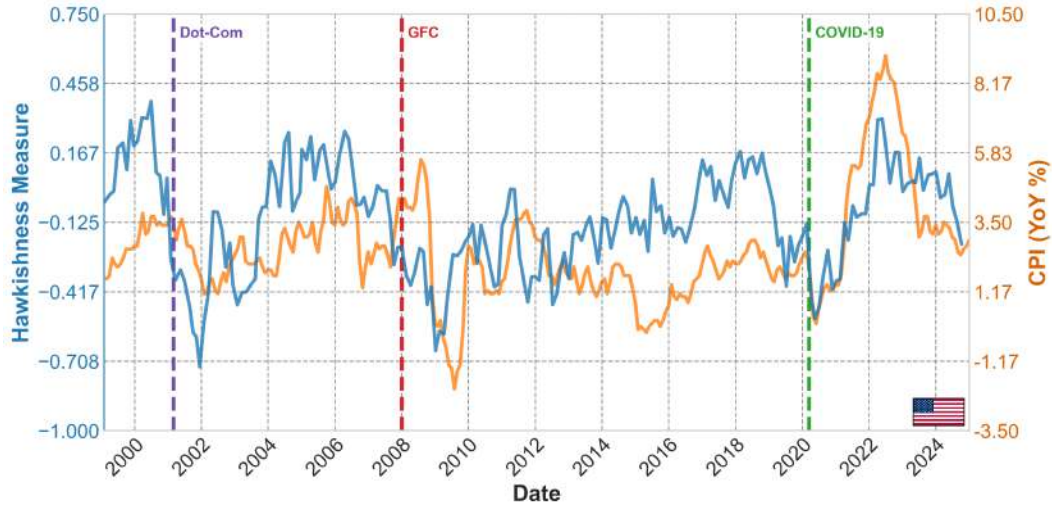


Figure 16: This figure illustrates how the events of the 2001 dot-com bust, 2008 financial crisis, and COVID-19 pandemic affect both the United States’s inflation index and our hawkishness measure. These periods were marked by dovish monetary policy as economic growth slowed and unemployment rose, and the FOMC cut rates to encourage spending. Additionally, notice how our hawkish measure aligns with the inflation index post-pandemic. Due to high stimulus spending, supply chain disruptions, and energy price shocks due to the events of Ukraine-Russia, prompted the FOMC to raise rates to bring inflation down to target levels.

1274 E.2 Hawkishness and Inflation

1275 We develop several dual-axis time-series plot for all banks individually. For all the plots, on the
 1276 left axis, we present model-derived hawkishness scores (defined in Section 6) from monetary-policy
 1277 communications whereas the right axis represents the specific economy’s inflation in terms of
 1278 Consumer Price Inflation (CPI).

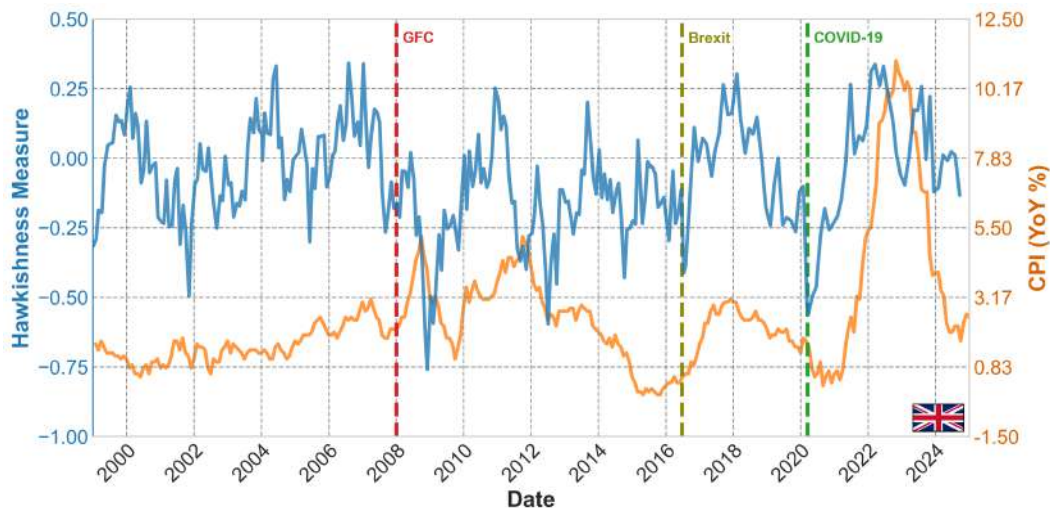


Figure 17: This figure illustrates how the events of the 2008 financial crisis and 2020 COVID-19 pandemic affect both the United Kingdom's inflation index and our hawkishness measure. These events were met with dovish monetary policies by BoE, such as asset buybacks and rate cutting, to stimulate the economy. Both measures also capture the uncertainty caused by Brexit in 2016 and high inflation rate in 2022 caused by supply chain disruptions and high consumer demand, which prompted the BoE to raise rates to stabilize the economy and lower inflation.

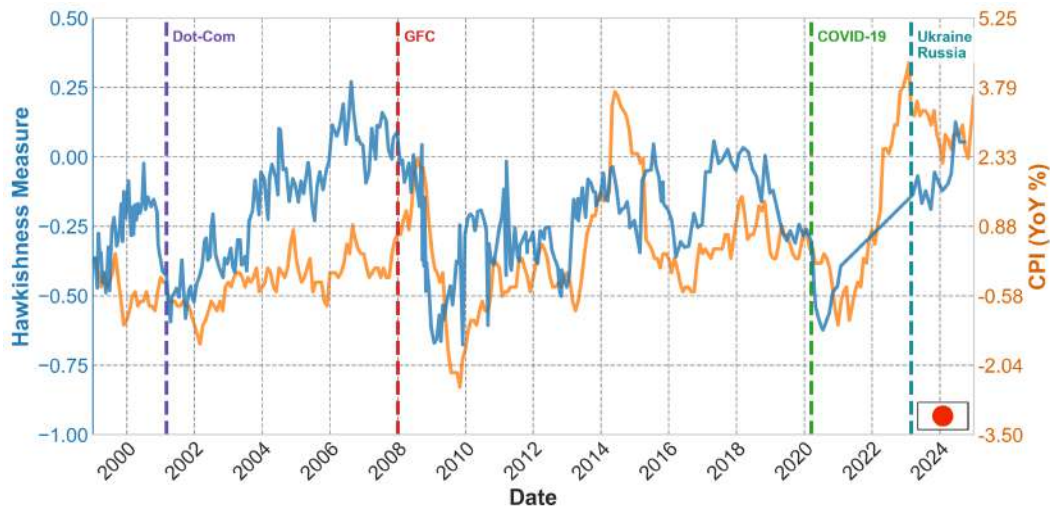


Figure 18: This figure illustrates how the events of the 2001 dot-com bust, 2008 financial crisis, and 2020 COVID-19 pandemic affect both the Japan's inflation index and our hawkishness measure. BoJ was extremely dovish to stimulate the economy during these times. Additionally, both metrics capture the high inflation and subsequent hawkish policy caused by supply chain disruptions due to the pandemic and rising commodity, energy, and food prices following the events of Ukraine-Russia.

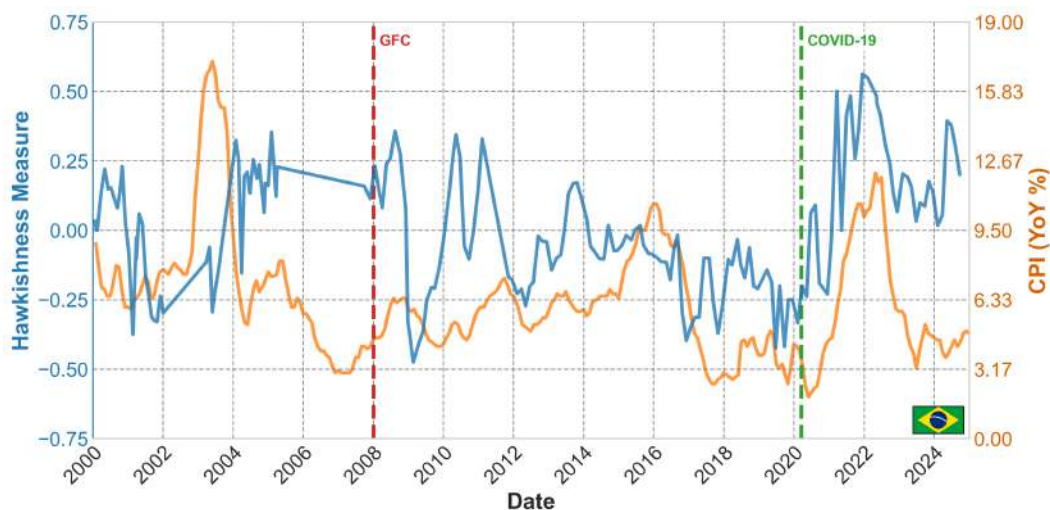


Figure 19: This figure illustrates how the events of the 2008 financial crisis as well as the COVID-19 pandemic affect both the Brazil's inflation index and our hawkishness measure. High levels of hawkishness during these periods reflects central bank's efforts to counteract rising inflation and return it to target levels. Additionally, the figure captures the heightened volatility in both inflation and hawkishness measures as the economy stabilizes following the 1999 financial crisis, during which high domestic debt was amounting to be 40% of GDP thus, leading to the devaluation of the Brazil real against the U.S. dollar.

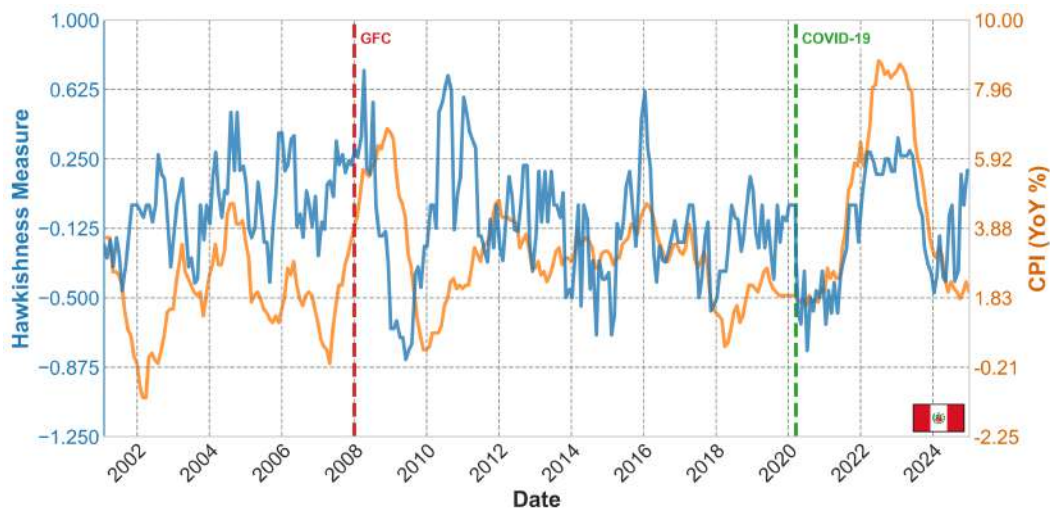


Figure 20: This figure illustrates how the events of the 2008 financial crisis and the 2020 COVID-19 pandemic affect both the Peru's inflation index and our hawkishness measure. BRCP implemented dovish monetary policy during these times to stimulate the economy. Both measures capture the surge in inflation and the corresponding spike in hawkishness caused by disruptions in supply chain and consumer demand, resulting in the BRCP raising rates to promote economic and price stability.

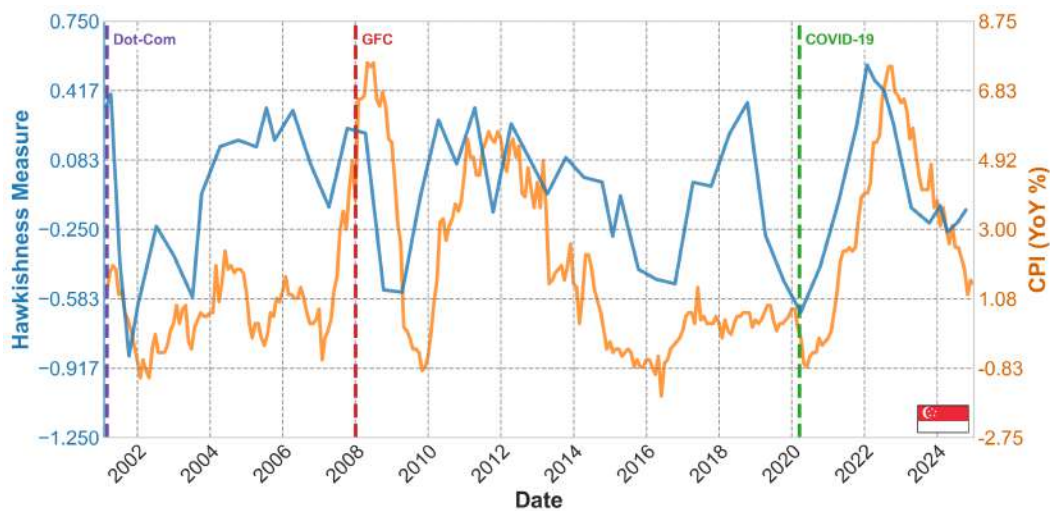


Figure 21: This figure illustrates how the events of the 2001 dot-com bust, 2008 financial crisis, and COVID-19 pandemic affect both the Singapore's inflation index and our hawkishness measure. These periods were marked by dovish monetary policy followed by high hawkishness as the economy contracted, and the MAS worked to stimulate the economy and control inflation after doing so.

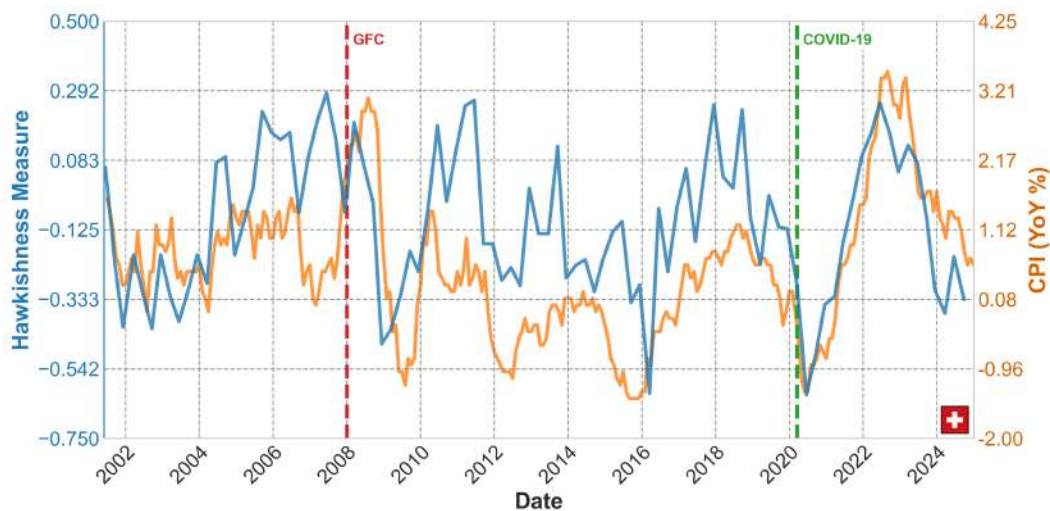


Figure 22: This figure illustrates how the events of the 2008 financial crisis and 2020 COVID-19 pandemic affect both the Switzerland's inflation index and our hawkishness measure. Both periods are marked by extremely low hawkishness due to the SNB trying to stimulate the economy. The figure also captures declining hawkishness in 2011 due to the Eurozone Crisis negatively impacting the exports. Our hawkishness measure also captures high inflation rates post-pandemic that SNB was combating.

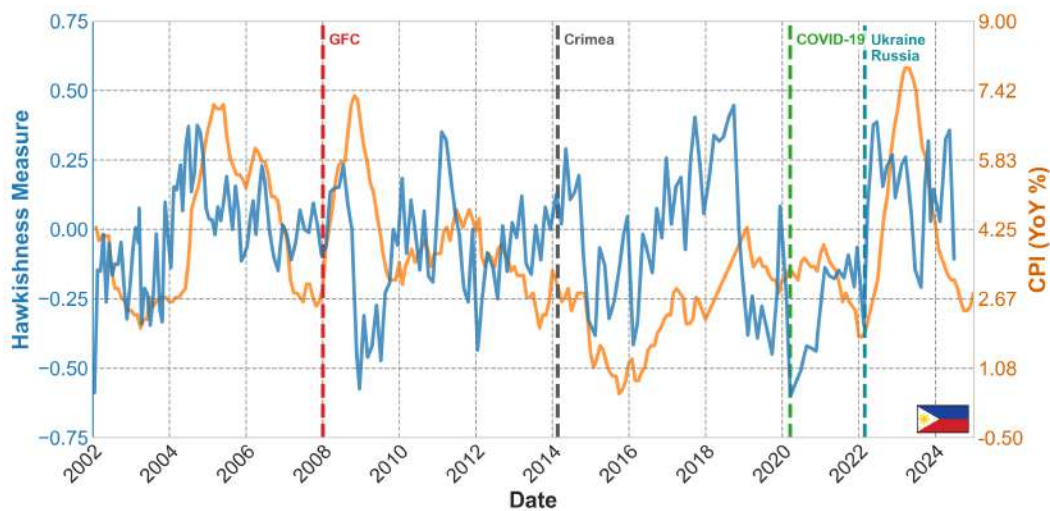


Figure 23: This figure illustrates how the events of the oil shocks in 2008 and 2014 affect both the Philippines's inflation index and our hawkishness measure. Following the oil shock in 2008, the CPI declines in response to increased hawkish monetary policy implementation. Additionally, the figure captures the effects of the 2020 COVID-19 pandemic within the inflation index and our hawkishness measure, resulting in dovish monetary policy to stimulate the economy. Both measures also capture the hawkish monetary policy due to surging food and energy prices from the events of Ukraine-Russia, prompting the BSP to raise rates in response to high inflation.

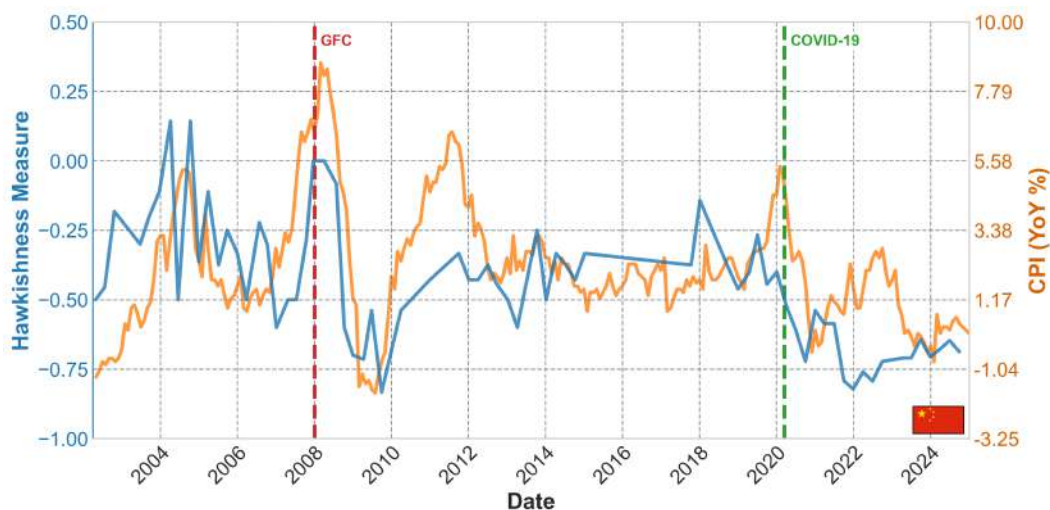


Figure 24: This figure illustrates how the events of the 2008 financial crisis and 2020 COVID-19 pandemic affect both the China's inflation index and our hawkishness measure. Both periods are marked by extremely low hawkishness due to the PBoC working to stimulate the economy.

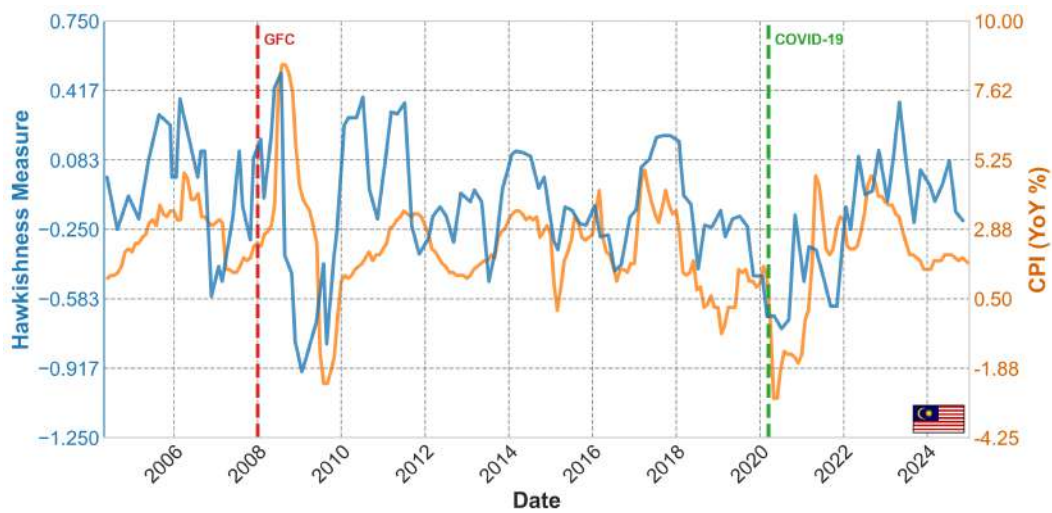


Figure 25: This figure illustrates how the events of the 2008 financial crisis and 2020 COVID-19 pandemic affect both the Malaysia's inflation index and our hawkishness. During these times BNM adopted dovish monetary policy to stimulate the economy. Both measures also capture the high inflation and subsequent hawkish policy caused by the strengthening of the US dollar, the events of Ukraine-Russia, and natural events which disrupted the agriculture sector and food chain.

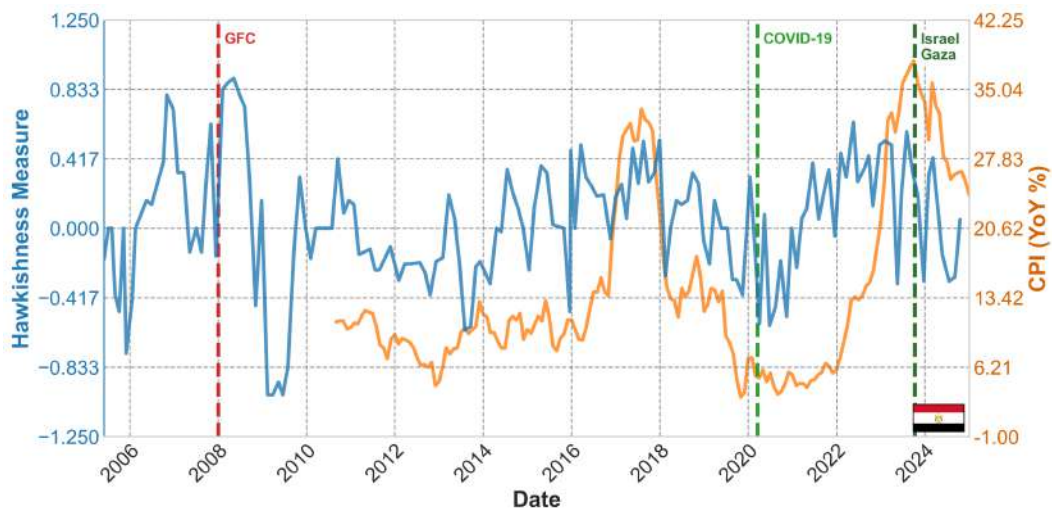


Figure 26: This figure illustrates the effects of the 2020 COVID-19 pandemic within the inflation index and our hawkishness measure during which dovish monetary policy prevailed in order to stimulate the economy. Ukraine-Russia hit the Egyptian economy hard, causing spike in oil and wheat prices and the depreciation of the EGP which led to rising inflation, prompting the CBE to raise rates multiple times.

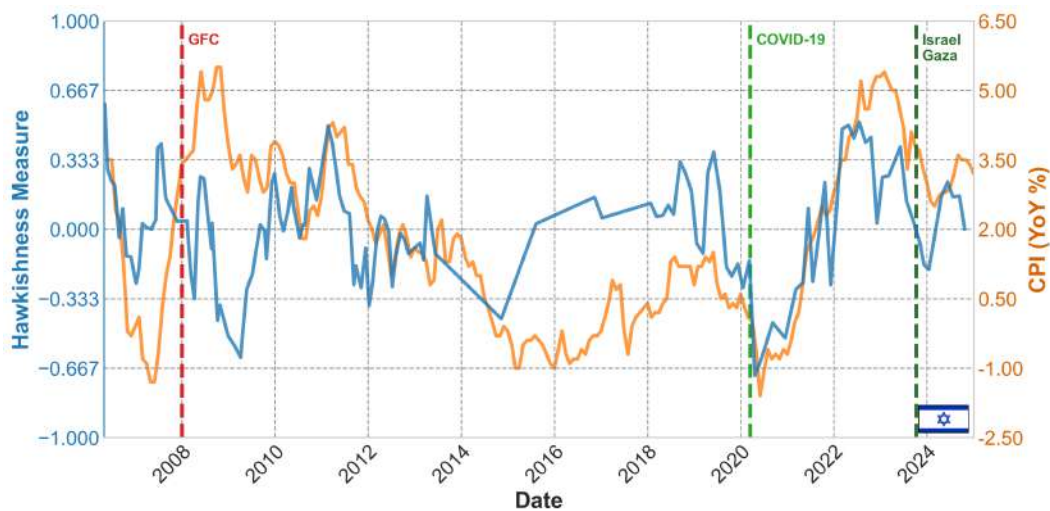


Figure 27: This figure illustrates how the events of the 2020 COVID-19 pandemic and the Israel-Hamas conflict in 2023 affect both the Israel's inflation index and our hawkishness measure. BoI implemented dovish monetary policy, including asset buybacks and rate cuts, to stimulate the economy during these times as major industries such as tourism were impacted.

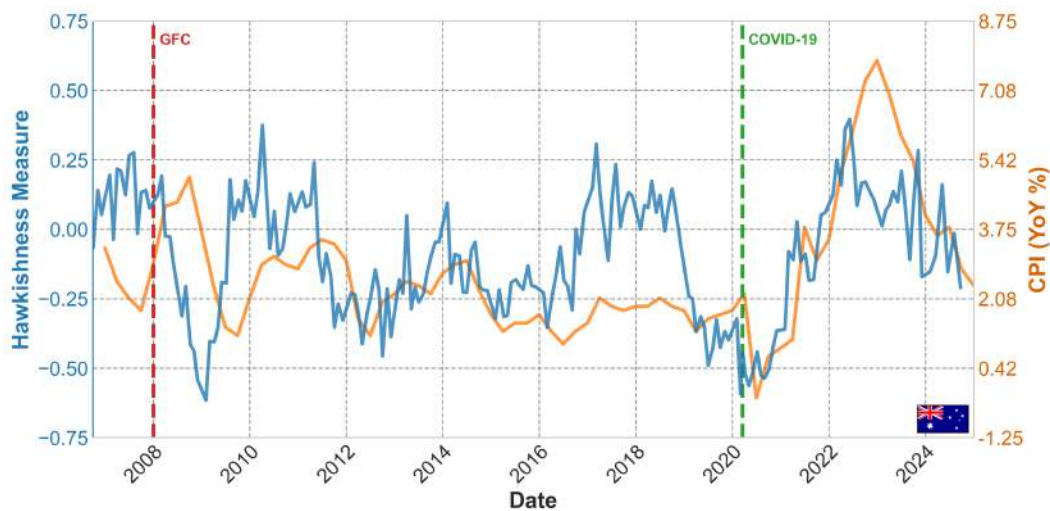


Figure 28: This figure illustrates how the events of the 2008 financial crisis and 2020 COVID-19 pandemic affect both the Australia's inflation index and our hawkishness measure. Both periods are marked by extremely low hawkishness due to the RBA working to stimulate the economy. During the pandemic, the efforts included lowering the cash rate, providing low-cost, long-term funding to banks, and purchasing government bonds. Our measure also captures the rising hawkishness post-pandemic, as the RBA worked to stabilize the economy.

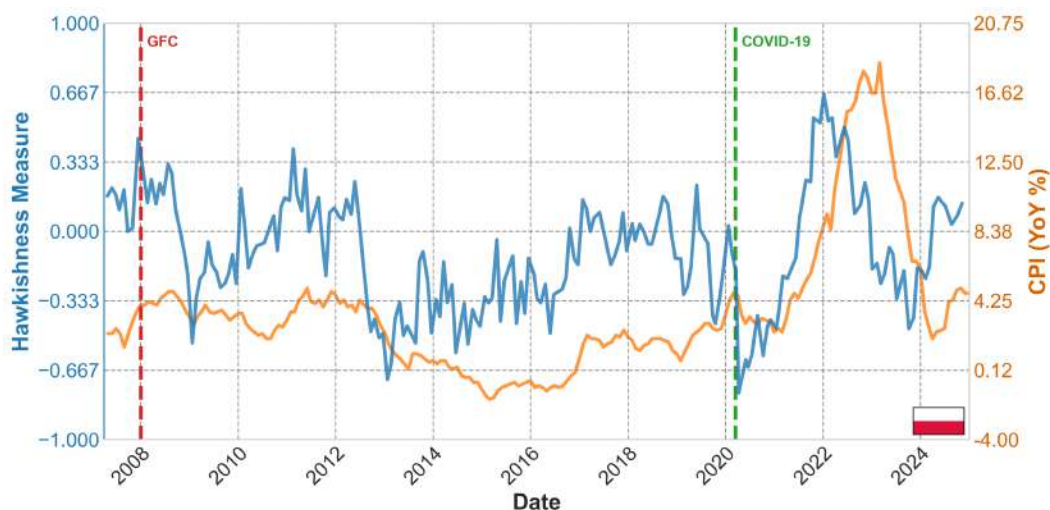


Figure 29: This figure illustrates how the events of the 2008 financial crisis and 2020 COVID-19 pandemic affect both the Poland's inflation index and our hawkishness measure. Both periods are marked by extremely low hawkishness due to the NBP working to stimulate the economy. Additionally, the figure captures hawkishness measure rising with the inflation index as the NBP raises rates to combat the high inflation level post-pandemic.

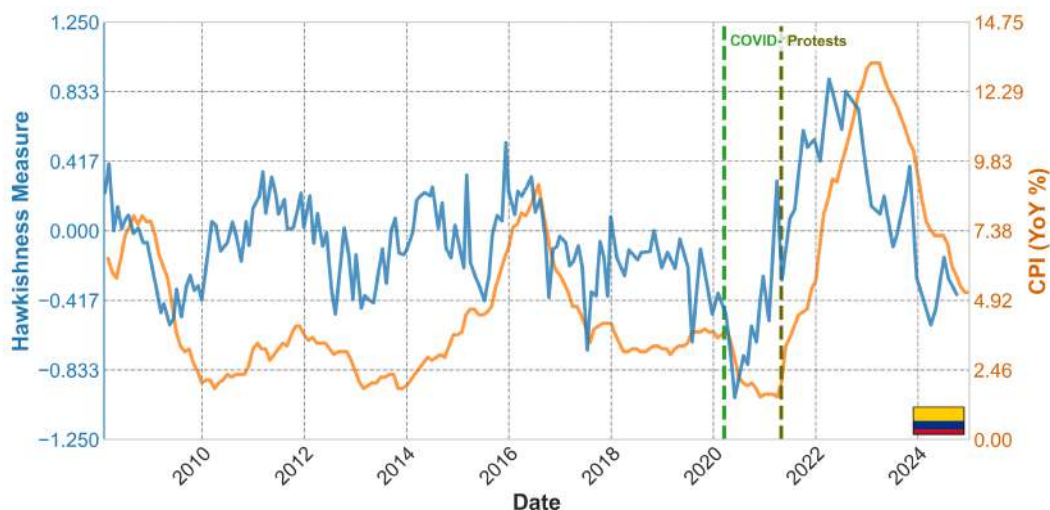


Figure 30: This figure illustrates how the events of the 2020 COVID-19 pandemic affect both the Colombia's inflation index and our hawkishness measure. BanRep implemented dovish monetary policy during these times to stimulate the economy. Both measures also capture the high inflation rate in 2022 caused by supply chain disruptions and high consumer demand, which prompted BanRep to raise rates to stabilize the economy and lower inflation.

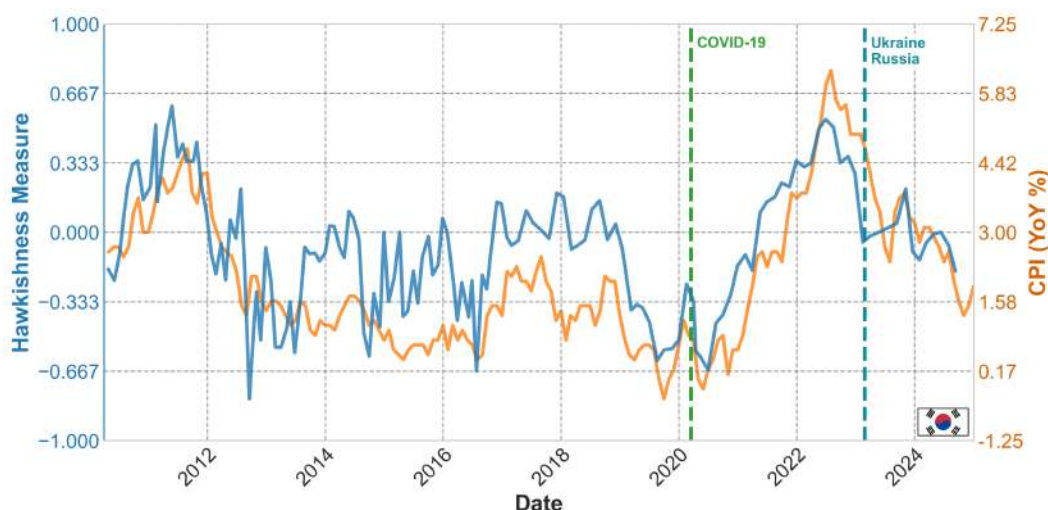


Figure 31: This figure illustrates the effects of the 2020 COVID-19 pandemic on South Korea's inflation index and our hawkishness measure. During this period, the dovish monetary policy was adopted by the central bank in order to stimulate the economy. Both indicators capture the steep incline in inflation in 2022, driven by the post-pandemic supply chain disruptions, increasing commodity prices, and strong consumer demand due to Ukraine-Russia. As a result, the BoK raised the interest rates to stabilize the economy and curb the inflation. Additionally, the effects of spillover from the Eurozone conflict is captured, which caused economic growth to slow, prompting dovish monetary policy.

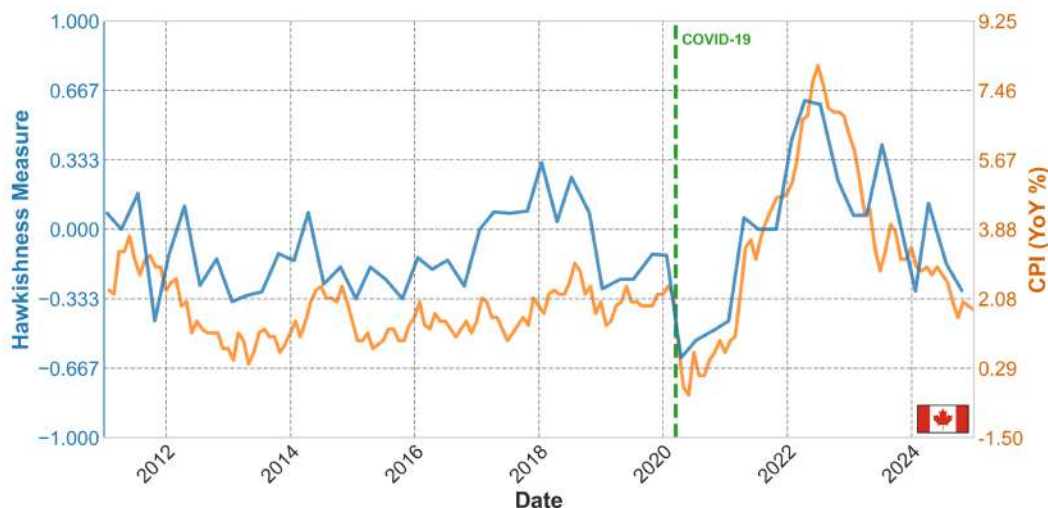


Figure 32: This figure illustrates the effects of the 2020 COVID-19 pandemic on the Canada's inflation index and our hawkishness measure. During this period, BoC implemented dovish monetary policy to stimulate the economy. Additionally, both measures capture the high inflation rate in 2022 caused by supply chain disruptions and high consumer demand, which prompted the BoC to raise rates to stabilize the economy and lower inflation.

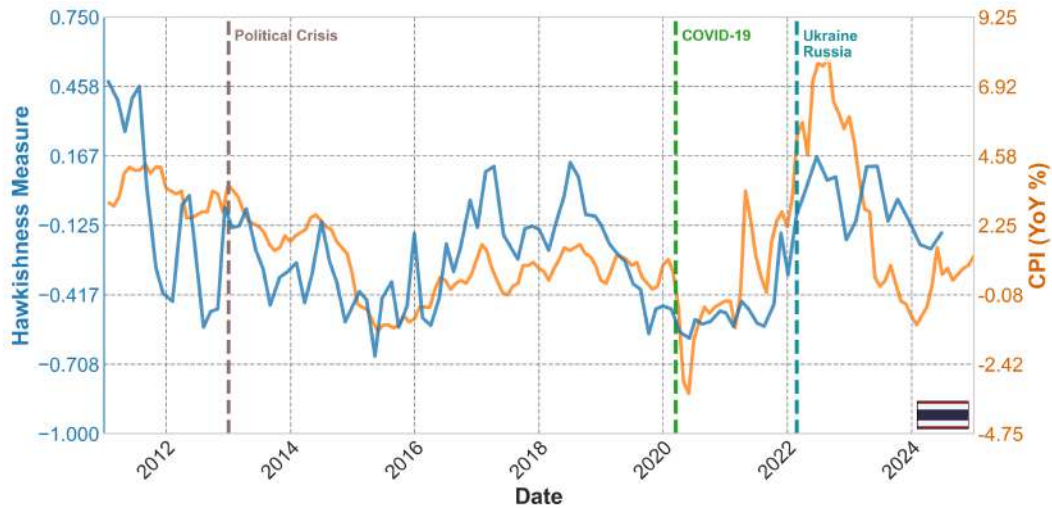


Figure 33: This figure illustrates the effects of the 2013-2014 Thai political crisis and COVID-19 pandemic on Thailand's inflation index and our hawkishness measure. During both periods, BoT adopted dovish monetary policies as the economy contracted. A recovery from the 2014 political unrest is captured by a rise in both the CPI and our hawkishness measure. Additionally, both measures reflect the increase in inflation and subsequent shift to hawkish policy in 2022 due to high energy and commodity prices caused by the events of Ukraine-Russia.

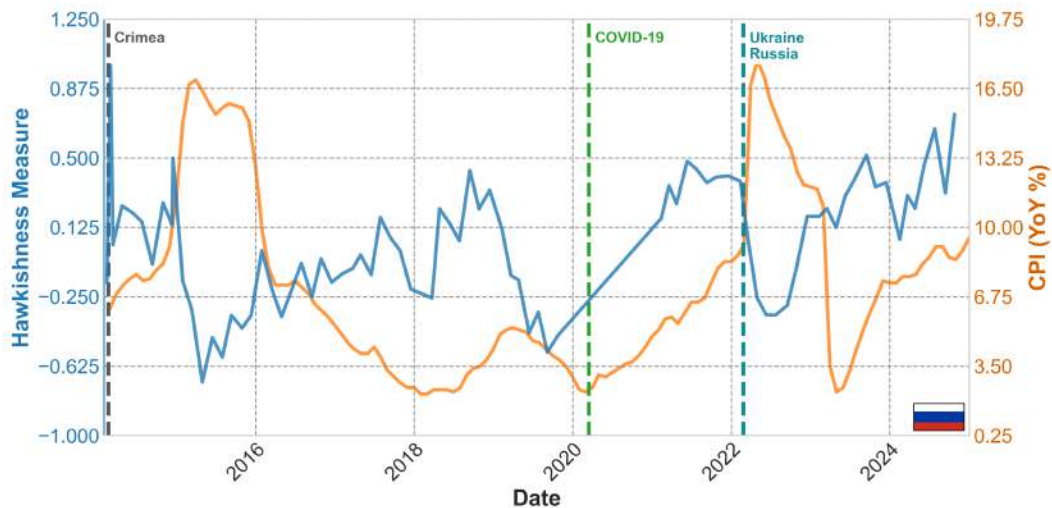


Figure 34: This figure illustrates the effects of major geopolitical and economic events on Russia's inflation index and our hawkishness measure. Elevated hawkishness is observed following the 2014 annexation of Crimea, reflecting CBR's efforts to curb the depreciation of the Russian ruble. The 2020 COVID-19 pandemic prompted dovish monetary policy (lowered interest rates) due to lowered GDP growth and CPI levels to stimulate the economy. The figure also captures the high inflation rate in late 2022 caused by supply chain disruptions due to the pandemic, and rising commodity prices and high consumer demand due to Ukraine-Russia, prompting the central bank to raise rates to stabilize the depreciating ruble.

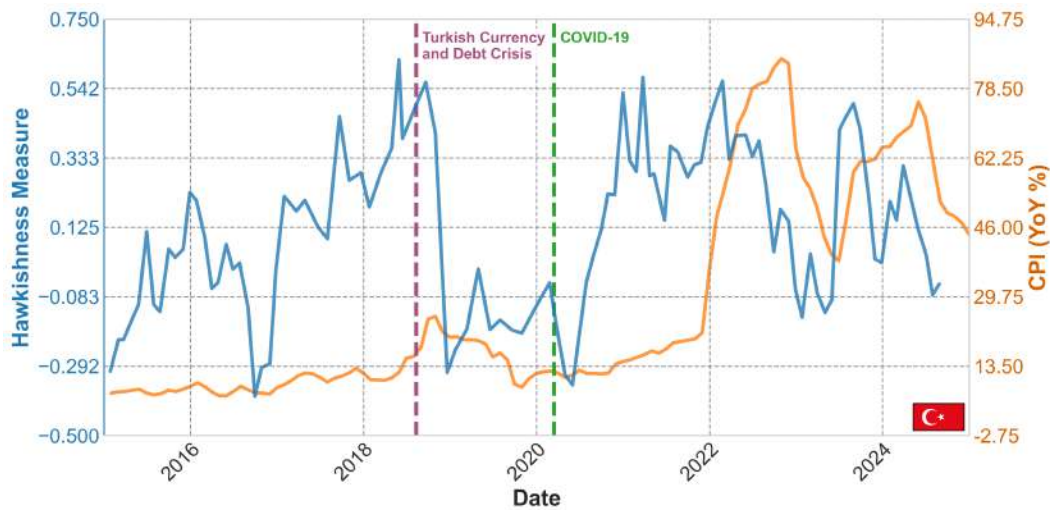


Figure 35: This figure illustrates the effects of the 2018-2019 economic crisis on the Turkey's inflation index and our hawkishness measure. The crisis sparked hawkish monetary policy to combat the Turkish lira depreciation, and then dovish monetary policy to stimulate economic growth to fund the deficit. Additionally, the figure captures the effects of the 2020 COVID-19 pandemic in both measures, and the high inflation rate in late 2022 caused by supply chain disruptions due to the pandemic, resulting in hawkish monetary policy to stabilize the Turkish lira.

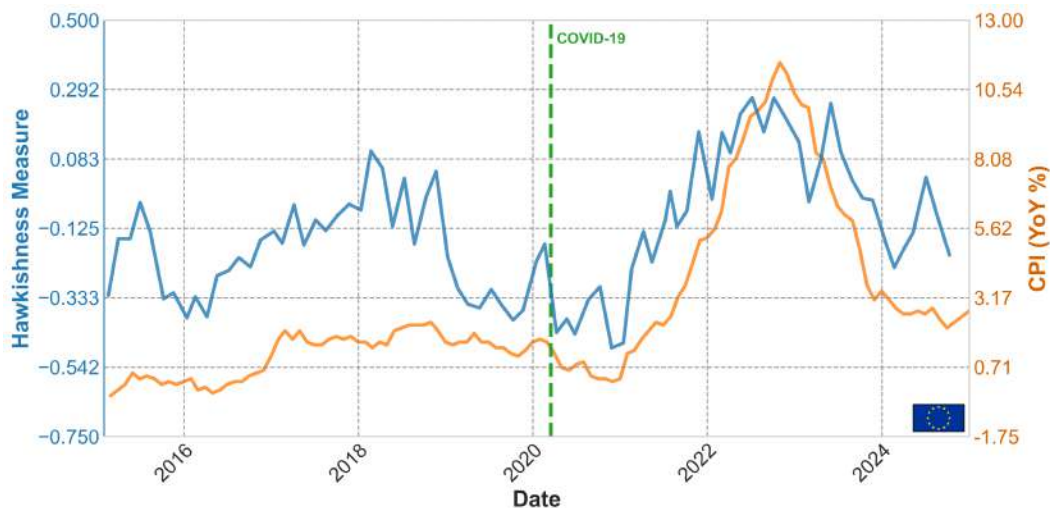


Figure 36: This figure illustrates the effects of the 2020 COVID-19 pandemic on the EU's inflation index and our hawkishness measure. During this period, dovish monetary policy prevailed as the ECB sought to stimulate the economy. Post-pandemic, our hawkishness measure also matches the inflation index when inflation sky-rocketed due to disrupted supply chains and rising demand and the ECB cut rates.

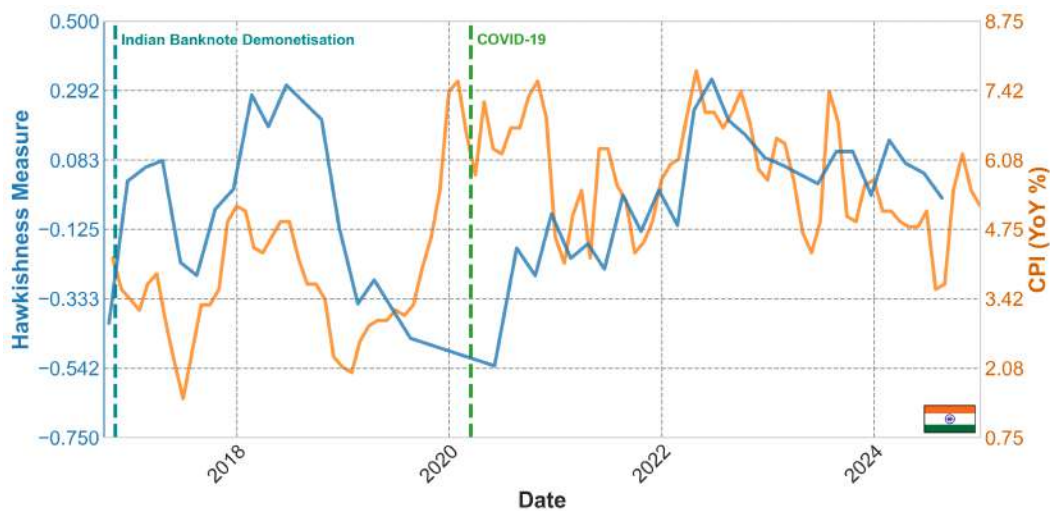


Figure 37: This figure illustrates the effects of the 2020 COVID-19 pandemic on the India's inflation index and our hawkishness measure. The period was marked by volatility, as the monetary policy aimed to support vulnerable households, encourage investment, and maintain price stability. The figure also captures the high hawkishness levels in 2018 reflecting the disinflationary measures implemented to combat liquidity issues following demonetization.

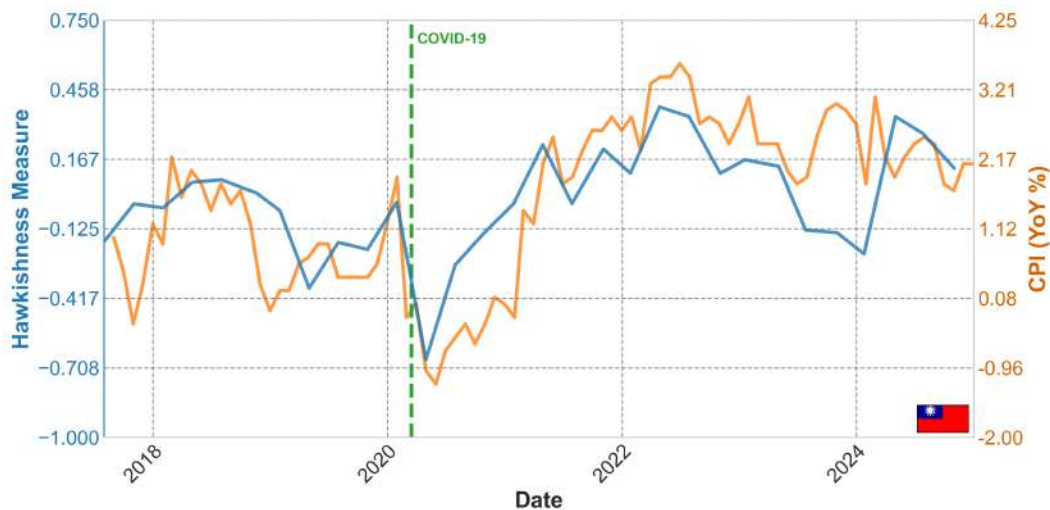


Figure 38: This figure illustrates the effects of the 2020 COVID-19 pandemic on the Taiwan's inflation index and our hawkishness measure. During this period, the central bank implemented dovish monetary policy to stimulate the economy. Additionally, the figure captures the high inflation rate in 2022 caused by supply chain disruptions and high consumer demand, which prompted the central bank to raise rates to stabilize the economy.

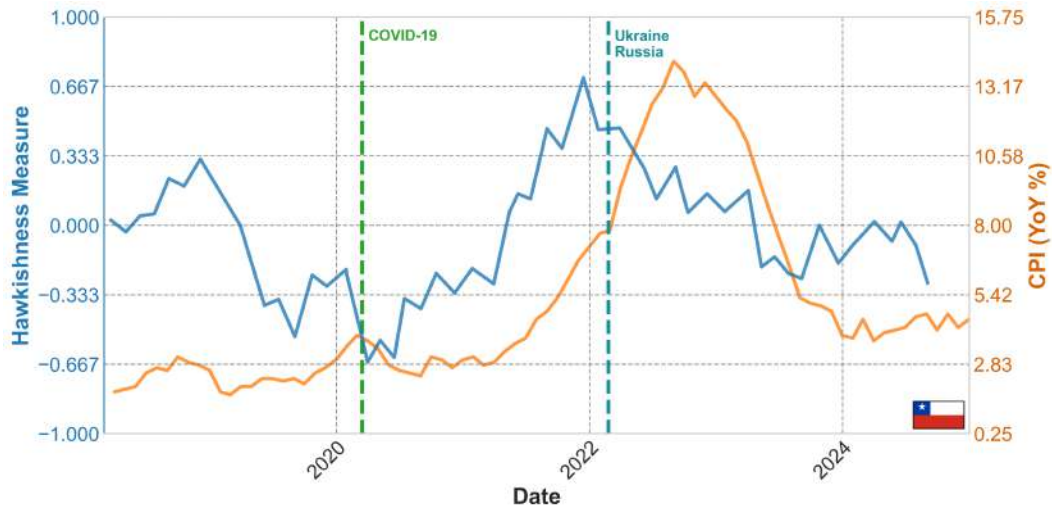


Figure 39: This figure illustrates how both the Chile's inflation index and our hawkishness measure capture the post-pandemic tightening of monetary policy. With a surge in demand coupled with supply-side bottlenecks, high commodity prices caused by Ukraine-Russia and depreciation of the Peso, the CBoC significantly raised policy rates.

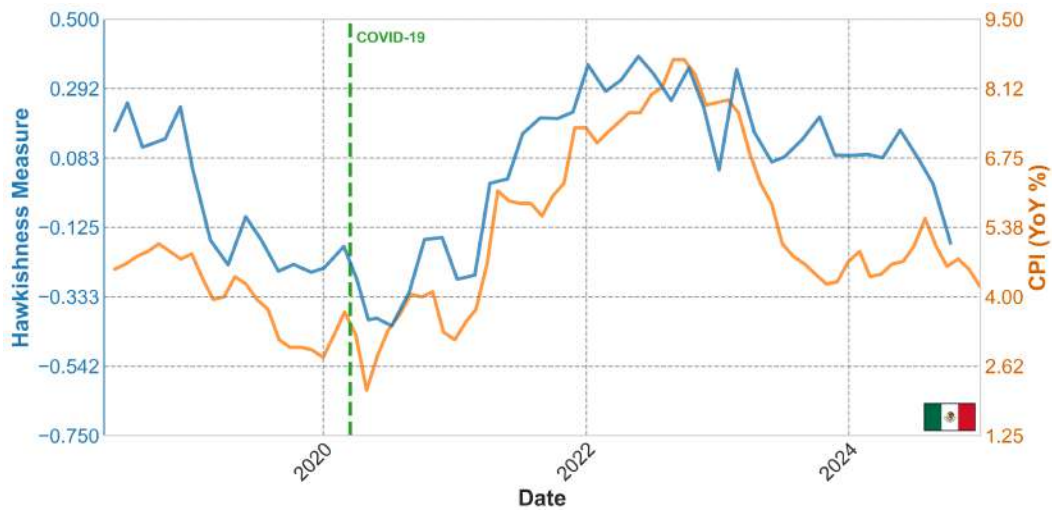



























Figure 40: This figure illustrates the effects of the 2020 COVID-19 pandemic on the Mexico's inflation index and our hawkishness measure. During this period, BoM implemented dovish monetary policy to stimulate the economy. Additionally, the figure captures the high inflation rate in 2022 caused by supply chain disruptions, rising commodity prices, and high consumer demand, which prompted the central bank to raise rates to stabilize the economy and lower inflation.

Table 29: Pearson correlation coefficients across all possible pairs of label categories (stance, temporal, certainty) for each bank over all years. Most label pairs exhibit a weak negative correlation. All values within brackets are negative.

	stance-temporal	stance-certainty	temporal-certainty
 BCB	0.003541	0.015641	(0.386366)
 BCRP	(0.060272)	0.108143	(0.414610)
 BNM	(0.063442)	0.010550	(0.507737)
 BSP	0.002618	0.028488	(0.538671)
 BanRep	0.019677	0.006071	(0.477708)
 BoC	(0.035004)	0.023413	(0.612648)
 BoE	(0.008751)	0.062980	(0.435042)
 BoI	(0.087278)	0.057606	(0.489864)
 BoJ	(0.074619)	0.026308	(0.388683)
 BoK	0.012551	0.010937	(0.390653)
 BdeM	0.083746	(0.026304)	(0.469378)
 BoT	(0.069446)	(0.032663)	(0.391205)
 CBCT	(0.040279)	0.063119	(0.488286)
 CBE	(0.095191)	0.077066	(0.464068)
 CBR	(0.029862)	(0.020771)	(0.386689)
 CBRT	(0.064327)	0.089533	(0.478669)
 CBoC	0.002561	(0.011903)	(0.305732)
 ECB	(0.016714)	(0.001914)	(0.351945)
 FOMC	(0.034457)	0.056154	(0.450257)
 MAS	(0.047880)	0.079492	(0.595583)
 NBP	0.005994	0.009087	(0.500110)
 PBoC	(0.283129)	(0.114589)	(0.007756)
 RBA	(0.000080)	0.035070	(0.465670)
 RBI	(0.017386)	0.034103	(0.401618)
 SNB	0.006077	0.010778	(0.452553)
Average	(0.035654)	0.023856	(0.434060)

To analyze the Temporal Classification and Uncertainty Estimation labels, we look at their correlation to see if these two labels are related. As seen in Table 29, we observe extremely low values of the Pearson correlation among the stance label with either of the two labels for any of the banks, indicating that Stance Detection is truly independent of both Temporal Classification and Uncertainty Estimation. We notice a weak negative correlation for every single central bank apart from the People's Bank of China. The weak negative correlation implies that there is a high probability that a Forward Looking statement is also Uncertain. Subsequently, Certain statements are likely to be Not Forward Looking.

F Extended Related Works

Research specifically focusing on Stance Detection spans multiple global contexts, including central banks from Israel [Kazinnik and Papell, 2021], Latin America (Colombia, Chile, Peru) [Vega and Lahura, 2020], Czech Republic, Hungary, Poland [Rozkrut et al., 2007], Japan [Oshima et al., 2018], Turkey [Iglesias et al., 2017], Mexico [Aguilar and Pérez-Cervantes, 2022], and India [Kumar et al., 2024]. While these studies provide valuable local insights, they typically lack comprehensive cross-bank and longitudinal analyses covering significant economic events [Ehrmann and Talmi, 2020, Wong and Fong, 2011, Armelius et al., 2020].

Our dataset uniquely addresses this gap by covering communications from 25 central banks between 1996 to 2024. It provides insights across Stance Detection, Uncertainty Estimation, and Temporal Classification labels, covering critical economic crises. Tables 2 and 30 compare our dataset comprehensively with other related datasets, highlighting our dataset’s unique breadth and temporal depth. Comparable datasets like MONOPOLY[Mathur et al., 2022] and EA-MD-QD[Barigozzi et al., 2024], among others, generally focus on narrower temporal ranges and fewer banks, excluding vital economic crisis periods covered in our analysis. Additionally, our experiments extends beyond benchmarking to human evaluation, Economic Analysis, error analysis, ablation studies on additional banks and on a Congress Committee hearing transcript.

Table 30: Comparison of existing datasets across year range and geographical coverage and inclusion of four specific additional analytical methods: error analysis, few shot benchmarking, predictive analysis, and systematic data collection. *Error Analysis* is defined as the evaluation of the reasoning of large language models on our specific tasks. *Few Shot Benchmark.* is defined as the inclusion of the task of few shot benchmarking for each of our LLMs and PLMs on our specific tasks. *Predictive Analysis* is defined as the process of using past data to predict future actions, specifically in the context of monetary policy meeting minutes and future policy implementations. *Systematic Data Collection* is defined as following a strict procedure during the collection of monetary policy data in a structured and methodical way.

Paper	Year Range	Geographical Range	Error Analysis	Few Shot Benchmark.	Predictive Analysis	Systematic Data Collection
Ivrendi and Yildirim [2013]	1995-2012	6 emerging economies	×	×	✓	×
Armelius et al. [2020]	2002-2017	23 central banks	×	×	✓	×
Mathur et al. [2022]	2009-2022	6 central banks	×	×	✓	✓
Kirti et al. [2022]	2020	74 countries	×	×	×	×
Shah et al. [2023]	1996-2022	United States	✓	×	✓	✓
Bolhuis [2024]	2000-2022	29 countries	×	×	✓	×
Barigozzi et al. [2024]	2000-2024*	Euro Area	×	×	✓	×
Zhang et al. [2025]	2009-2024	United States	✓	×	×	✓
WCB (Ours)	1996-2024	25 central banks	✓	✓	✓	✓

1305 G Glossary

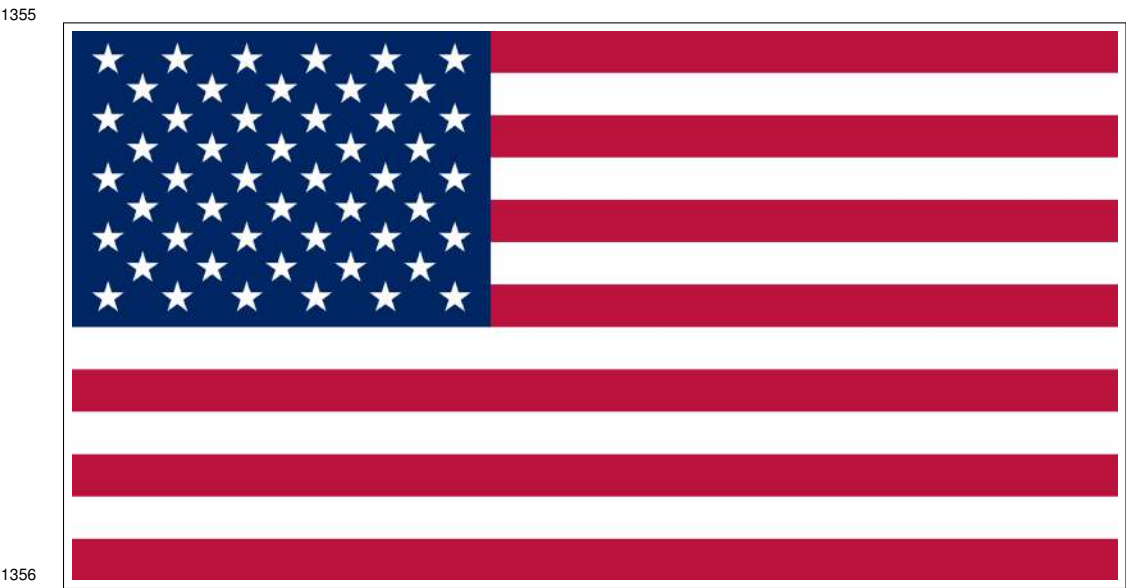
1306 The definitions of common financial terms used in this paper are as follows:

- 1307 • **Central Bank:** A central bank is a financial institution responsible for the formulation of
1308 monetary policy and the regulation of member banks. It typically has privileged control
1309 over the production and distribution of money and credit for a nation or a group of nations.
1310 [Segal et al., 2023]
- 1311 • **Certainty:** The state where the outcome of a decision or policy is known or predictable
1312 with a high degree of confidence.
- 1313 • **Consumer Price Index (CPI):** CPI is calculated as a weighted average of prices for selected
1314 items, with weights reflecting their share in total consumer spending. It is a statistical
1315 measure that tracks the average change over time in the prices paid by households for a
1316 representative basket of goods and services, and is widely used as a measure for inflation.
1317 [Fernando et al., 2025, Hayes et al., 2024]
- 1318 • **Dovish Policy:** A type of policy that favors lower interest rates and quantitative easing to
1319 encourage economic growth. [Shah et al., 2023]
- 1320 • **Forward Looking:** A business term used to identify predictions that corporations and
1321 institutions make about future business or economic conditions.
- 1322 • **Harmonized Index of Consumer Prices (HICP):** The Harmonized Index of Consumer
1323 Prices (HICP) provides the official measure of consumer price index in the euro area and the
1324 EU. [Eurostat, 2024]
- 1325 • **Hawkish Policy:** A type of policy that favors higher interest rates and quantitative tightening
1326 to keep inflation in check. [Shah et al., 2023]
- 1327 • **Inflation:** Inflation is a gradual loss of purchasing power that is reflected in a broad rise in
1328 prices for goods and services over time. [Fernando et al., 2024]
- 1329 • **Interest Rate:** Interest rate is the amount charged on top of the principal by a lender to a
1330 borrower for use of assets, expressed as a percentage or ratio. Central banks use the interest
1331 rate as a monetary policy tool. Increasing the cost of borrowing among commercial banks
1332 can influence many other interest rates (like mortgage rates, personal loan rates), making
1333 it difficult to borrow money. This leads to lower demand for money and can cool the hot
1334 economy and vice versa. [Banton et al., 2025]
- 1335 • **Irrelevant:** In this paper's context, it refers to a statement in our dataset that does not convey
1336 any indication of the central bank's monetary policy stance.
- 1337 • **Monetary Policy:** Monetary policy is a set of tools used by a nation's central bank to
1338 control the overall money supply and promote economic growth, employing strategies such
1339 as revising interest rates and changing bank reserve requirements. [Segal et al., 2023]
- 1340 • **Neutral:** In this paper's context, it refers to a neutral monetary policy that does not favor
1341 either hawkishness or dovishness. [Shah et al., 2023]
- 1342 • **Not Forward Looking:** It refers to statements or information focusing on past events or
1343 historical facts rather than speculating about the future.
- 1344 • **Not Seasonally Adjusted (NSA):** NSA means that the economic data (e.g., GDP) have
1345 not been adjusted to remove the effects of predictable seasonal fluctuations. [Majaski et al.,
1346 2022]
- 1347 • **Purchasing Power:** Purchasing power refers to the amount of goods and services one unit
1348 of currency can buy. [Hayes et al., 2024]
- 1349 • **Uncertainty:** The state where there is doubt regarding a particular outcome and a possibility
1350 of different outcomes.

1351 **H Central Banks Information, Annotator Disagreement, and Annotation**
1352 **Guides**

1353 **Federal Open Market Committee**

1354 ‘ **Region: United States of America**



1356

1357 **Data Collected: 1996-2024**

1358

1359

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

1360 Monetary Policy Mandate

1361 The FOMC is responsible for formulating U.S. monetary policy to achieve its dual mandate: promot-
1362 ing maximum employment and ensuring price stability.

1363 Mandate Objectives:

- 1364 • **Maximum Employment:** Striving for the highest *sustainable* level of employment, recog-
1365 nizing that various factors influence employment levels.
- 1366 • **Price Stability:** Aiming for a 2% annual inflation rate, as measured by the Personal
1367 Consumption Expenditures (PCE) Price Index, to maintain purchasing power and economic
1368 stability.

1369 Structure

1370 FOMC is integral to the Federal Reserve System, overseeing U.S. monetary policy.

1371 Composition:

- 1372 • **Board of Governors:** Seven members appointed by the President and confirmed by the
1373 Senate, serving staggered 14-year terms.
- 1374 • **Federal Reserve Bank Presidents:**
 - 1375 – Permanent Member: President of the Federal Reserve Bank of New York.
 - 1376 – Rotating Members: Four of the remaining eleven Reserve Bank presidents, serving
1377 one-year terms on a rotating basis.

1378 Meeting Structure:

- 1379 • **Frequency:** Eight scheduled meetings annually, approximately every six weeks.
- 1380 • **Additional Meetings:** Held as needed to address urgent economic developments.

1381 Manual Annotation

1382 Annotators

- 1383 • Liam Dolphin
- 1384 • Tiberius Colina
- 1385 • Pranav Aluru
- 1386 • Ahaan Limaye

1387 Annotation Agreement

1388 The agreement percentage among the pairs of annotators for different labels.

- 1389 • Stance Agreement: 47.8%
- 1390 • Time Agreement: 82.5%
- 1391 • Certainty Agreement: 75.1%

1392 Annotation Guide

1393 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
1394 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
1395 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
1396 other than monetary policy stance.

1397 The annotation guide for monetary policy stance is summarized in Table 31. It is built by dividing each
1398 target sentence into eight defined categories: Economic Status, Dollar Value Change, Energy/House
1399 Prices, Foreign Nations, Fed Expectations/Actions/Assets, Money Supply, Key Words/Phrases,
1400 corporate bond, and Labor.

- 1401 • *Economic Status*: A sentence pertaining to the state of the economy, relating to unemploy-
1402 ment and inflation.
- 1403 • *Dollar Value Change*: A sentence pertaining to changes such as appreciation or depreciation
1404 of value of the United States Dollar on the Foreign Exchange Market.
- 1405 • *Energy/House Prices*: A sentence pertaining to changes in prices of real estate, energy
1406 commodities, or energy sector as a whole.
- 1407 • *Foreign Nations*: A sentence pertaining to trade relations between the United States and a
1408 foreign economy. If not discussing the United States we label neutral.
- 1409 • *Fed Expectations/Actions/Assets*: A sentence that discusses changes in the Fed yields, bond
1410 value, reserves, or any other financial asset value.
- 1411 • *Money Supply*: A sentence that overtly discusses impact to the money supply or changes in
1412 demand.
- 1413 • *Key Words/Phrases*: A sentence that contains key word or phrase that would classify it
1414 squarely into one of the three label classes, based upon its frequent usage and meaning
1415 among particular label classes.
- 1416 • *Corporate Bond*: A sentence that relates to corporate bond issuance.
- 1417 • *Labor*: A sentence that relates to changes in labor productivity.

1418 **Examples:**

- 1419 • “The Committee then turned to a discussion of the economic and financial outlook, the
1420 ranges for the growth of money and debt in 1996, and the implementation of monetary
1421 policy over the intermeeting period ahead.”
1422 **Neutral**: Economic status is only mentioned without diving deeper into increases or de-
1423 creases in inflation and/or unemployment rate.
- 1424 • “To support the Committee’s decision to raise the target range for the federal funds rate, the
1425 Board of Governors voted unanimously to raise the interest rates on required and excess
1426 reserve balances to 2.”
1427 **Hawkish**: The Fed expects higher inflation so they are likely to increase interest rates.
- 1428 • “Labor productivity has continued to rise over recent months, supporting stronger economic
1429 growth and reducing inflationary pressures.”
1430 **Dovish**: The increase in labor productivity can ease inflation concerns without requiring
1431 tighter monetary policy.
- 1432 • “The U.S. trade deficit has widened significantly due to increased imports from foreign
1433 nations, contributing to upward pressure on inflation.”
1434 **Hawkish**: A growing trade deficit can fuel inflation, prompting the Fed to consider tightening
1435 monetary policy to control price stability.
- 1436 • “Real GDP was anticipated to increase at a rate noticeably below its potential in 2008.”
1437 **Dovish**: Expected GDP below potential indicates that economy is slowing down, requiring
1438 expansionary monetary policy to prevent any drawbacks.
- 1439 • “It was agreed that the next meeting of the committee would be held on Tuesday-Wednesday,
1440 April 26-27, 2016.”
1441 **Irrelevant**: This sentence solely provides scheduling information and does not offer any
1442 insights into a monetary policy stance.

Table 31: The Federal Open Market Committee Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Economic Status	When inflation increases, when unemployment decreases, when economic growth is projected high, when economic output is higher than potential supply/actual output, when economic slack falls.	When inflation decreases, when unemployment increases, when economic growth is projected as low.	When unemployment rate or growth is unchanged, maintained, or sustained.	Sentence is not relevant to monetary policy.
Dollar Value Change	When the dollar depreciates.	When the dollar appreciates.	N/A	Sentence is not relevant to monetary policy.
Energy/House Prices	When oil/energy prices increase, when house prices increase.	When oil/energy prices decrease, when house prices decrease.	N/A	Sentence is not relevant to monetary policy.
Foreign Nations	When the US trade deficit increases.	When the US trade deficit decreases.	When relating to a foreign nation's economic or trade policy.	Sentence is not relevant to monetary policy.
Fed Expectations, Actions, and Assets	Fed expects high inflation, widening spreads of treasury bonds, increase in treasury security yields, increase in TIPS value, increase bank reserves.	Fed expects subpar inflation, Fed expecting disinflation, narrowing spreads of treasury bonds, decreases in treasury security yields, and reduction of bank reserves.	N/A	Sentence is not relevant to monetary policy.
Money Supply	Money supply is high, increased demand for goods, low demand for loans.	Money supply is low, M2 increases, increased demand for loans.	N/A	Sentence is not relevant to monetary policy.
Key Words and Phrases	Indicating a focus on "price stability" and "sustained growth."	When the stance is "accommodative," indicating a focus on "maximum employment" and "price stability."	Use of phrases "mixed," "moderate," "reaffirmed."	Sentence is not relevant to monetary policy.
Corporate Bond	When issuance increases.	When issuance decreases.	N/A	Sentence is not relevant to monetary policy.
Labor	When productivity or unemployment decreases.	When productivity or unemployment increases.	N/A	Sentence is not relevant to monetary policy.

1443 The annotation guide for whether the sentence has Forward Looking information or not is summa-
1444 rized in the Table 32 and whether the committee is talking about the sentence with Certainty attest
1445 to it or not in Table 33.

Table 32: The Federal Open Market Committee Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	When it discusses expectations, projections, or anticipations about future economic conditions, policy actions, or resource utilization.	“On the other hand, given the high level of capacity utilization and movements in the exchange rate, there is a danger that higher production costs will increasingly be passed on to prices.”
Not Forward Looking	When it reflects on recent or past economic data, trends, or events to describe what has already occurred.	“By contrast, exports of goods and services stagnated.”

Table 33: The Federal Open Market Committee Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	When an expectation, trend, action, or outcome is definitively stated without ambiguity. Key words include “remain,” “increased,” “decreased,” and “will continue.”	“This is because the utilization of resources is somewhat higher.”
Uncertain	When it expresses concerns, potential risks, or need to alter monetary policy if a specific event comes to pass. Key words include “outlook,” “if,” “could,” and “uncertainty.”	“These higher prices for imports could impact upon consumer prices.”

1446

People’s Bank of China

1447

Region: China

1448



1449

1450

Data Collected: 2003-2024

1451

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

1452

1453 Monetary Policy Mandate

1454 The People’s Bank of China (PBoC) is responsible for determining and implementing the monetary
1455 policy of China upon the guidance and advice of the Monetary Policy Committee (MPC). The PBoC
1456 also maintains the circulation of the Renminbi and monitors monetary money-laundering suspicions,
1457 and also “promot[es] the building up of [the] credit information system”.

1458 Mandate Objectives:

- 1459 • **Currency Stability:** Maintain stability of the value of the Renminbi to promote economic
1460 growth.

1461 Structure

1462 Composition:

- 1463 • **Governor**
 - 1464 – Nominated by the Premier of the State Council and decided by the National People’s
1465 Congress.
 - 1466 – When the National People’s Congress is not in session, the Governor is decided by the
1467 Standing Committee of the National People’s Congress and appointed by the President
1468 of the People’s Republic of China.
- 1469 • **Deputy Governors:** There are a number of deputy governors. As of March 2025, there are
1470 five.

1471 **Meeting Structure:** It is unclear when the PBoC meets. However, the MCB, the consultancy body,
1472 meets as follows:

- 1473 • **Frequency:** Meetings are on a quarterly basis
- 1474 • **Additional Meetings:** Held if one-third of the members agree or if proposed by the
1475 Chairman

1476 Manual Annotation

1477 Annotators

- 1478 • Henry Zhang
- 1479 • Sidharth Subbarao
- 1480 • Harold Huang
- 1481 • Rachel Yuh

1482 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 1483 • Stance Agreement: 63.2%
- 1484 • Time Agreement: 90.7%
- 1485 • Certainty Agreement: 91.5%

1486 Annotation Guide

1487 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
1488 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
1489 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
1490 other than monetary policy stance.

1491 The annotation guide for monetary policy stance is summarized in Table 34. It is built by dividing
1492 each target sentence into seven defined categories: Inflation, Value of the Yuan, Interest Rates, Money
1493 Supply, GDP Growth Rate, and Trade Output.

- 1494 • *Inflation*: A sentence about the inflation rate in the economy
- 1495 • *Value of the Yuan*: A sentence about how the value of the Yuan is changing relative to the
- 1496 USD
- 1497 • *Interest Rates*: A sentence about how interest rates, including the repo rate and loan prime
- 1498 rates are changing.
- 1499 • *Money Supply*: A sentence pertaining the M2 money supply. PCB may conduct short-term
- 1500 liquidity operations and reverse repo operations to influence liquidity into the market.
- 1501 • *GDP Growth Rate*: A sentence pertaining to changes in GDP or changes in the commodities
- 1502 market or key industry that may impact the GDP
- 1503 • *Trade Output*: A sentence pertaining to China's export or import levels or how macroeco-
- 1504 nomic factors may impact foreign trade

1505 **Examples:**

- 1506 • "In line with the rules on monetary policy committee of the people's bank of china and with
- 1507 the approval of the state council, presidents of two wholly state-owned commercial banks,
- 1508 Liu Mingkang, Zhang Enzhao and Wu Jinglian will withdraw from the monetary policy
- 1509 committee as their term has expired."
- 1510 **Irrelevant:** Sentence is not directly related to monetary policy.
- 1511 • "Improvement of the managed floating exchange rate regime should be effected, continuously
- 1512 following the principle of "making it a self-initiated, controllable and gradual process.""
- 1513 **Dovish:** The PBoC is confident in the health of the market is not concerned about the
- 1514 money supply becoming too large hence their desire to focus on floating the exchange rate.
- 1515 • "It was noted at the meeting that efforts should made to implement the prudent monetary
- 1516 policy, make the financial measures more targeted, flexible and effective, and give more
- 1517 priority to stabilizing the general price level in 2011."
- 1518 **Hawkish:** The PBoC feels that monetary policy needs to be implemented in order to
- 1519 stabilize prices, which is dovish.
- 1520 • "The committee discussed monetary policy and measures to be adopted in the period ahead
- 1521 and concurred that the sound monetary policy should be preserved, interest rates of deposits
- 1522 and loans be maintained stable to ensure steady advancement of market-based interest rate
- 1523 reform."
- 1524 **Neutral:** The PBoC is emphasizing maintaining the currency policy.

Table 34: People's Bank of China Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Inflation	When inflation rates rise inflationary pressures.	When inflation rates fall.	When inflation rate is stable.	Sentence is not relevant to monetary policy.
Value of the Yuan	When Yuan depreciates.	When Yuan appreciates.	When Yuan remains stable.	Sentence is not relevant to monetary policy.
Interest Rates	When the repo rate increases or loan prime rates (LPR) are increased.	When repo rate decreases or when short-term rates are lowered.	When interest rates are maintained at the same level.	Sentence is not relevant to monetary policy.
Money Supply	When M2 money supply is decreasing.	When PBC restricts money supply.	When money supply remains the same.	Sentence is not relevant to monetary policy.
GDP Growth Rate	When GDP growth declines.	When GDP growth increases or real economic growth is promoted.	When GDP growth remains constant.	Sentence is not relevant to monetary policy.
Trade Output	When policy aims to decrease foreign trade.	When there are bottlenecks key industries namely oil, coal, and electricity, or sluggish global economy, or other hindrances to foreign trade.	When there is a focus on maintaining stability in trade.	Sentence is not relevant to monetary policy.

1525 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 35 and whether the committee is talking about the sentence with Certainty attest
1526 to it or not in Table 36.
1527

Table 35: People’s Bank of China Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	When it discusses future outcomes, key words include “in the future,” “expected,” and “projected.”	“The credit information system and capital raising system for some should be established.”
Not Forward Looking	When it references past events, historical data, previous trends, or observed outcomes, and may include key words such as “was,” “in the past,” and “has happened.”	“Liquidity was adequate in the banking system, growth of money and credit supply rapid and the financial system operated in a stable manner.”

Table 36: People’s Bank of China Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	When it include indicators based on reliable data and trends, expresses outcomes definitively, includes phrases such as “will,” “definitely,” and “certainly.”	“A mix of new monetary policy tools will be used to keep liquidity adequate at a reasonable level.”
Uncertain	When it mentions metrics with high volatility such as global commodity prices or includes qualifying clauses that include hedging words such as “maybe,” “kind of,” “sort of,” “might,” or “possibility.”	“While financial support to economic growth must be guaranteed, no negligence should be tolerated in preventing inflation and financial risks.”

1528

1529

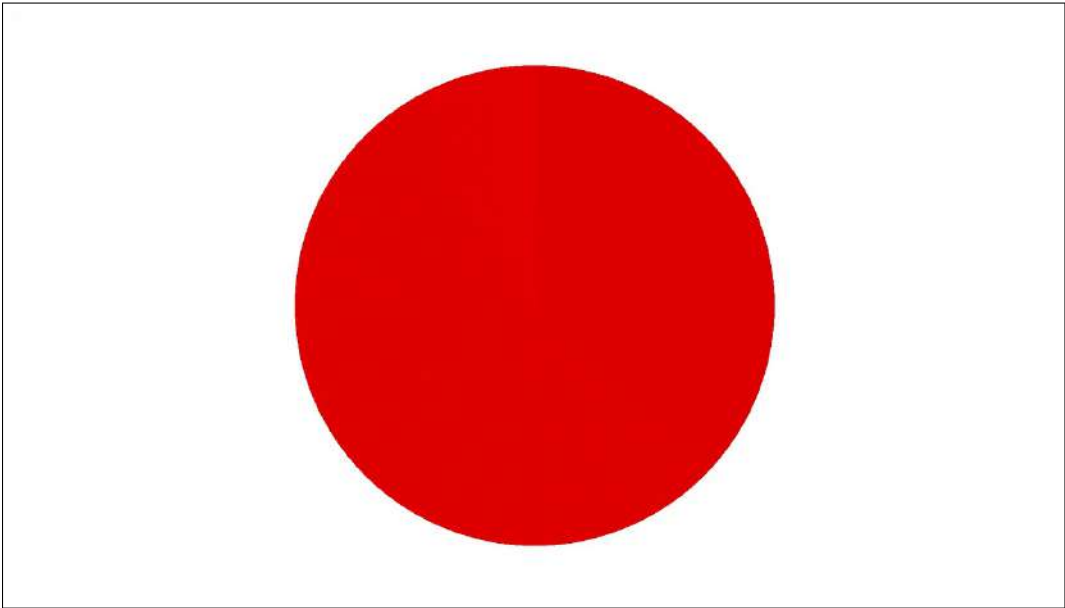
1530

Bank of Japan

Region: Japan

1531

1532



Data Collected: 1998-2024

1533

1534

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

1535 **Monetary Policy Mandate**

1536 The Bank of Japan implements monetary policy to maintain price stability, which underpins economic
1537 activity. It influences interest rates through money market operations to control currency and liquidity.
1538 The Policy Board sets the policy direction at Monetary Policy Meetings (MPMs), where economic
1539 and financial conditions are reviewed. Decisions are announced immediately, guiding daily market
1540 operations to provide or absorb funds as needed.

1541 **Mandate Objectives:**

- 1542 • **Issuance of Banknotes and Monetary Policy Implementation:** The Bank of Japan is
1543 responsible for issuing banknotes and conducting currency and monetary policy operations
1544 to ensure effective monetary control.
- 1545 • **Financial Stability:** The Bank strives to facilitate the smooth settlement of funds among
1546 banks and other financial institutions, thereby contributing to the overall stability of the
1547 financial system.
- 1548 • **Economic Development:** Through its currency and monetary policy measures, the Bank
1549 aims to achieve price stability, fostering the sound and sustainable development of the
1550 national economy. The current inflation target is set at 2%.

1551 **Structure**

1552 The bank's structure consists of the policy board and the bank's officers (Governor, Deputy Governors,
1553 Auditors, Executive Directors, and Counsellors.).

1554 **Composition:**

- 1555 • **Policy Board:** The Policy Board is established as the Bank's highest decision-making
1556 body. The Board determines the guideline for currency and monetary control, sets the basic
1557 principles for carrying out the Bank's operations, and oversees the fulfillment of the duties
1558 of the Bank's officers, excluding Auditors and Counselors.
- 1559 • **Bank's Officers:** The number of posts given to the Bank's officers is as follows: the
1560 Governor, two Deputy Governors, six Members of the Policy Board, at most three Auditors,
1561 at most six Executive Directors, and the Counselors.

1562 **Meeting Structure:**

- 1563 • **Frequency:** Eight scheduled meetings annually, approximately every six weeks.
- 1564 • **Additional Meetings:** Held as needed to address urgent economic developments.

1565 **Manual Annotation**

1566 **Annotators**

- 1567 • Randall Bevan
- 1568 • Alex Chen
- 1569 • Abinav Chari
- 1570 • Aksh Tyagi

1571 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 1572 • Stance Agreement: 37%
- 1573 • Time Agreement: 62.6%
- 1574 • Certainty Agreement: 71.3%

1575 **Annotation Guide** Our annotation guide is inspired by the work of Shah et al. [2023]. We modify
 1576 their annotation guide further, keeping in mind the changing global macroeconomic dynamics. Also,
 1577 we have two additional categories (Forward Looking or Not Forward Looking and Certain or
 1578 Uncertain) of labels other than monetary policy stance.

1579 The annotation guide for monetary policy stance is summarized in Table 37. It is built by dividing each
 1580 target sentence into eleven defined categories: Money Supply, Economic Status, Yen Value Change,
 1581 Energy and Commodity Prices, Foreign Trade, Monetary Policy Stance, Inflation Expectations,
 1582 Government Bonds and JGB Yields, BoJ Asset Purchases, Labor Market Conditions, and Fiscal
 1583 Policy Interactions.

- 1584 • *Money Supply*: A sentence that discusses changes in the central bank's control over the
 1585 money supply of Japan. This includes adjustments in interest rates, quantitative easing
 1586 policies, and shifts in inflation tolerance, which for Japan, is currently at 2%. Japan's current
 1587 quantitative easing policy involves buying \$40 billion worth of Japanese government bonds
 1588 monthly.
- 1589 • *Economic Status*: A sentence that describes the overall condition of the economy, highlight-
 1590 ing trends in inflation, consumer spending, and economic growth.
- 1591 • *Yen Value Change*: A sentence pertaining to fluctuations in the value of the Yen on the
 1592 foreign exchange market.
- 1593 • *Energy and Commodity Prices*: A sentence addressing movements in energy and commodity
 1594 markets or addressing policy that is expected to cause change in energy and commodity
 1595 markets. For example, lower commodity prices in 2010s allowed the Bank of Japan to keep
 1596 policies loose without inflation fears.
- 1597 • *Foreign Trade*: A sentence that focuses on foreign trade (e.g., imports and exports) and its
 1598 impacts on Japan's economy. For example, if Japan has a trade deficit, money would be, on
 1599 net, leaving the Japanese economy.
- 1600 • *Monetary Policy Stance (Interest Rates and actions)*: A sentence that outlines the central
 1601 bank's policy direction through interest rate decisions and asset purchase programs, signaling
 1602 either a dovish or hawkish stance.
- 1603 • *Inflation Expectations*: A sentence mentions inflation expectations in Japan. Japan's current
 1604 inflation target is at 2%.
- 1605 • *Government Bonds and JGB Yields*: A sentence discussing the behavior of government bond
 1606 yields, including new adjustments in government bond purchases or the yield curve.
- 1607 • *BoJ Asset Purchases (ETFs, REITs, etc.)*: A sentence related to the Bank of Japan's strategy
 1608 in buying financial assets such as ETFs.
- 1609 • *Labor Market Conditions*: A sentence focused on the labor market, including indicators
 1610 such as unemployment rates and wage growth.
- 1611 • *Fiscal Policy Interactions*: A sentence that examines the relationships between monetary and
 1612 fiscal policies, highlighting how changes in, for example, government spending or taxation
 1613 affect monetary policy.

1614 **Examples:**

- 1615 • "Some members commented that the growth in demand for some it-related goods was
 1616 slowing around the world."
 1617 **Dovish**: The subsection in the sentence about slowing demand implies that current economic
 1618 pressures are subdued, increasing the need for expansionary monetary policy.
- 1619 • "With regard to the first factor, one member said that inflation expectations for fiscal 2012
 1620 and 2013 had been relatively low at around 0 percent, although medium- to long-term
 1621 inflation expectations remained stable at around 1.0 percent."
 1622 **Dovish**: The emphasis on very low near-term inflation expectations would allow for expan-
 1623 sionary monetary policy without inflation fears.
- 1624 • "Some members said that private consumption and fixed asset investment continued to show
 1625 high growth."
 1626 **Hawkish**: The noted high growth in private consumption and fixed asset investment indicate

1627 an overheating economy, which may eventually generate inflationary pressures. This would
 1628 ultimately require contractionary monetary policy to offset any future inflationary pressures
 1629 that may arise.

- 1630 • "Stock prices in emerging economies had also dropped significantly, due mainly to the rise
 1631 in inflation."
 1632 **Hawkish:** The sentence indicates that inflationary pressures are affecting asset markets,
 1633 requiring a tightening stance to counteract the inflation.
- 1634 • "The funding conditions of European financial institutions had recently been even more
 1635 stable."
 1636 **Neutral:** The statement objectively reports that funding conditions are stable without
 1637 implying any need for either easing or tightening monetary policy.
- 1638 • "The representative from the cabinet office made the following remarks."
 1639 **Irrelevant:** This sentence is purely introductory and does not provide any substantive
 1640 economic or policy-related information for analysis.

Table 37: Bank of Japan Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Money Supply	Lower interest rates, quantitative easing via bank purchases, higher inflation tolerance.	Higher interest rates on money, cease quantitative easing (QE), lower inflation tolerance. A 0.25% interest rate is the highest in Japan currently.	If the interest rates are staying the same.	Sentence is not relevant to monetary policy.
Economic Status	Liquid economy, increased money flow between businesses and consumers, increasing CPI.	More rigid economy that is shrinking from a state of inflation, less exchange of money between businesses and consumers, decreasing CPI.	If the CPI stays the same or is not affected.	Sentence is not relevant to monetary policy.
Yen Value Change	When the Yen appreciates (Ex. Allowing the yen to appreciate to reduce import costs).	When the Yen depreciates (Ex. Allowing the yen to weaken to boost exports).	Maintaining a stable Yen exchange rate.	Sentence is not relevant to monetary policy.
Energy and Commodity Prices	When energy and commodity prices decrease, thus allowing the BoJ to continue their easing measures.	When energy and commodity prices increase.	Stable energy and commodity prices lead the BoJ to make no significant changes.	Sentence is not relevant to monetary policy.
Foreign Trade	An increase in trade and exports.	A decrease in trade or exports to reduce inflation.	Foreign trade and exports are maintained.	Sentence is not relevant to monetary policy.
Monetary Policy Stance (Interest Rates and actions)	Expanding asset purchases or lowering interest rates (Ex. Lowering interest rates and purchasing assets to support the economy, such actions took place during COVID-19).	Reducing asset purchases or signaling future interest rate increases (Ex. BoJ reducing its target for JGB purchases or signaling a future interest rate increase to curb inflation).	Keeping monetary policy unchanged, no signs of making interest rate changes or modifying practices of purchasing assets.	Sentence is not relevant to monetary policy.
Inflation Expectations	High or increasing inflation.	Low or decreasing inflation (target 2%).	If inflation is unchanged.	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
Government Bonds and JGB Yields	When BoJ increases bond purchases or targets a 10-year JGB yield near 0% and indicates readiness to lower the yield further.	When BoJ allows a rise in 10-year JGB yields above 1% by reducing bond purchases or adjusting yield curve control target further.	10-year JGB yields remain within the target range (currently $0\% \pm 1\%$) with no significant change.	Sentence is not relevant to monetary policy.
BoJ Asset Purchases (ETFs, REITs, etc.)	When BoJ increases asset purchases, injecting liquidity into the market.	When BoJ reduces asset purchases.	No change in the BoJ asset purchase strategy.	Sentence is not relevant to monetary policy.
Labor Market Conditions	When unemployment rate is high or is increasing or wage growth is low or decreasing, indicating a need for accommodative policy.	When unemployment rate is low or is decreasing, or wage inflation is high or is increasing.	When unemployment rate remains between 2.5% - 3.5% with annual wage growth around 1.5% - 2.5%.	Sentence is not relevant to monetary policy.
Fiscal Policy Interactions	Correlated with expansionary fiscal policies (Ex. BoJ gave government stimulus during crisis situations such as 2008 economic crisis and COVID-19 pandemic).	Correlated with contractionary fiscal policies (tax increases, reduced spending).	No connected changes with fiscal policy.	Sentence is not relevant to monetary policy.

1641 The annotation guide for whether the sentence has Forward Looking information or not is summa-
1642 rized in the Table 38 and whether the committee is talking about the sentence with Certainty attest
1643 to it or not in Table 39.

Table 38: Bank of Japan Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	When the text mainly uses future tense. Key words include “will,” and “is expected to.”	“The Japanese economy was recovering at a moderate pace and was expected to continue to do so.”
Not Forward Looking	When the text is using past or present tense. Key words include “has,” and “did.”	“This member continued, however, that these changes did not reflect an economic downturn.”

Table 39: Bank of Japan Annotation Guide to identify whether a sentence echoes certainty or not.

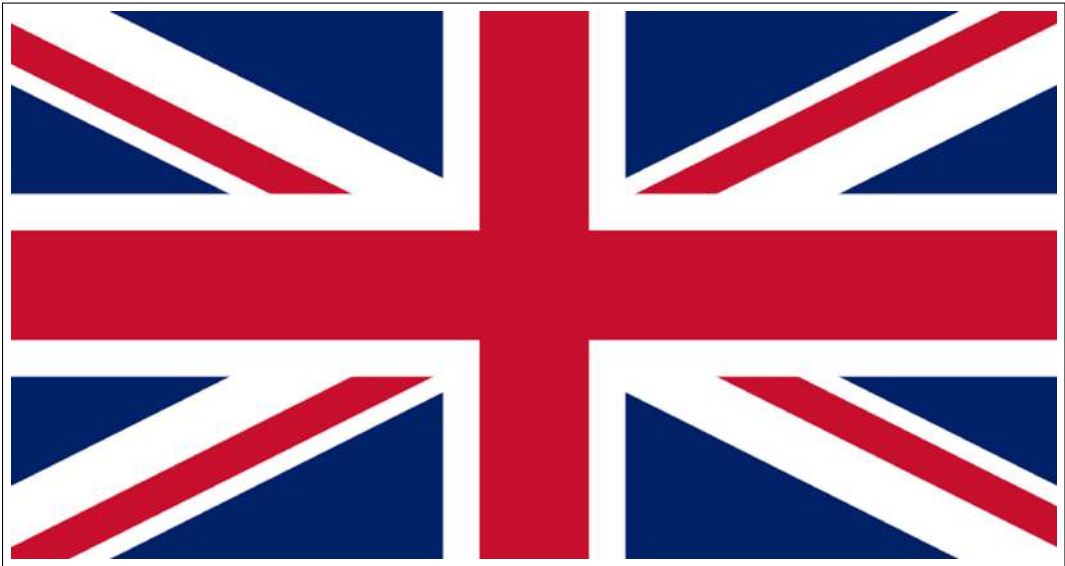
Label	Description	Example
Certain	When the text uses firm language. Key words include “will,” “increased,” “decreased.”	“The Japanese economy was on the way to recovery at a moderate pace.”
Uncertain	When the text uses cautious language. Key words include “may” and “could.”	“Exports were expected to continue to rise against the background of the expansion of overseas economies as a whole.”

1644
1645
1646

1647

1648

Bank of England
Region: United Kingdom



Data Collected: 1999-2024

1649

1650

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

1651 Monetary Policy Mandate

1652 The Bank of England sets monetary policy to maintain low and stable inflation in the UK.

1653 Mandate Objectives:

- 1654 • **Stable Inflation:** The primary objective of monetary policy is to ensure low and stable
1655 inflation. This is achieved by targeting a 2% inflation rate over the medium term, as mandated
1656 by the Government.
- 1657 • **Sustainable Economic Growth:** In addition to inflation control, the Bank supports the
1658 Government's broader economic objective of fostering strong, sustainable, and balanced
1659 economic growth.

1660 Structure

1661 Monetary policy decisions are made by the Monetary Policy Committee (MPC), which consists of
1662 nine members. Each member serves a fixed term, after which they may be reappointed or replaced.

1663 Composition:

- 1664 • **Monetary Policy Committee (MPC):** The MPC comprises nine members: the Governor,
1665 three Deputy Governors (responsible for Monetary Policy, Financial Stability, and Markets
1666 & Banking), the Chief Economist, and four external members appointed by the Chancellor.
- 1667 • **External Members:** These members are appointed to provide independent perspectives and
1668 bring expertise from outside the Bank of England.
- 1669 • **HM Treasury Representative:** A representative of HM Treasury attends MPC meetings,
1670 but does not have voting rights. Their role is to provide insights on fiscal policy and other
1671 government economic initiatives while ensuring the Chancellor remains informed about
1672 monetary policy decisions.

1673 **Meeting Structure:** The MPC sets and announces policy eight times a year (roughly once every six
1674 weeks). In the week leading up to each announcement, the committee meets several times.

- 1675 • **Pre-MPC Meeting:** Members receive a briefing from staff on the latest economic data and
1676 trends. This includes insights on business conditions across the UK from regional agents.
- 1677 • **First Meeting:** Members review and discuss the most recent economic indicators. This
1678 meeting typically takes place in the week leading up to the policy announcement.
- 1679 • **Second Meeting:** Members deliberate on potential monetary policy actions. This discussion
1680 occurs a few days before the official announcement.
- 1681 • **Final Meeting:** The Governor proposes a policy stance expected to gain majority support,
1682 followed by a vote among MPC members. This meeting usually takes place on Wednesday,
1683 one day before the decision is announced. Each member has a vote, and the policy with the
1684 most votes is selected. In the case of a tie, the Governor casts the deciding vote. Members
1685 in the minority express their preferred policy stance.
- 1686 • **MPC Announcement:** The final decision, along with meeting minutes, is published at
1687 12 noon on the day of the announcement, typically a Thursday. Rate announcements are
1688 disseminated both on the Bank's website and directly to market participants.

1689 Manual Annotation

1690 Annotators

- 1691 • Austin Barton
- 1692 • Srimaaye Jegannathan
- 1693 • Andrew Li
- 1694 • Aidan Wu

1695 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 1696 • Stance Agreement: 59.3%
- 1697 • Time Agreement: 77.4%
- 1698 • Certainty Agreement: 68.5%

1699 **Annotation Guide**

1700 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
1701 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
1702 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
1703 other than monetary policy stance.

1704 The annotation guide for monetary policy stance is summarized in Table 40. It is built by dividing each
1705 target sentence into eight defined categories: Economic Status, Pound Value Change, Energy/House
1706 Prices, Foreign Nations, BoE Expectations, Actions, Assets, Money Supply, Key Words, Phrases,
1707 and Labor.

- 1708 • *Economic Status*: A sentence pertaining to the overall condition of the economy, including
1709 indicators such as inflation, unemployment, and growth projections.
- 1710 • *Pound Value Change*: A sentence pertaining to fluctuations in the value of the British Pound
1711 on the foreign exchange market, indicating whether it is appreciating or depreciating.
- 1712 • *Energy/House Prices*: A sentence pertaining to variations in the prices of energy commo-
1713 dities and residential properties.
- 1714 • *Foreign Nations*: A sentence pertaining to the economic impacts of trade relations with
1715 foreign countries, particularly those affecting the United Kingdom.
- 1716 • *BoE Expectations, Actions, Assets*: A sentence that discusses the Bank of England’s pro-
1717 jections, policy actions, and changes in financial assets such as treasury bonds and bank
1718 reserves.
- 1719 • *Money Supply*: A sentence that examines changes in the overall availability of money in the
1720 economy, including shifts in loan demand and monetary liquidity.
- 1721 • *Key Words, Phrases*: A sentence that includes specific keywords or phrases which clearly
1722 indicate a particular monetary policy stance.
- 1723 • *Labor*: A sentence that relates to shifts in labor market conditions, encompassing changes in
1724 productivity, wage trends, and job openings.

1725 **Examples:**

- 1726 • “Recent economic indicators showing rising inflation and falling unemployment have led
1727 the Bank of England to signal a tighter monetary policy stance.”
1728 **Hawkish**: The BOE is concerned that overheating growth will fuel further inflation, prompt-
1729 ing interest rate hikes.
- 1730 • “The acceleration of house prices and energy costs is causing upward pressures on consumer
1731 demand, which may necessitate a proactive tightening of monetary policy.”
1732 **Hawkish**: Rising asset prices and energy costs are seen as signals of mounting inflationary
1733 pressures.
- 1734 • “Falling inflation paired with a rise in unemployment has convinced policymakers at the
1735 Bank of England that a more accommodative stance is needed to support the economy.”
1736 **Dovish**: Lower inflation and softer labor market conditions suggest that further easing could
1737 help sustain economic growth.
- 1738 • “Low money supply growth combined with persistently high loan demand has reinforced
1739 the BOE’s outlook for maintaining an expansionary monetary policy stance.”
1740 **Dovish**: These factors indicate subdued economic activity, reducing the likelihood of
1741 immediate monetary policy tightening.
- 1742 • “Total output prices had been flat year on year.”
1743 **Neutral**: The sentence states that total output prices remained stable and doesn’t take a
1744 direction on monetary policy.

1745
1746
1747

- “The new government had announced its intention to present an updated budget on 8 July.”
Irrelevant: The announcement pertains to a general government schedule and does not contain a monetary policy stance.

Table 40: Bank of England Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Economic Status	Inflation decreases, unemployment increases, low projected growth.	Inflation increases, unemployment decreases, high projected growth, demand outgrows supply.	When unemployment rate or growth is unchanged, maintained, or sustained.	Sentence is not relevant to monetary policy.
Pound Value Change	Pound appreciates.	Pound depreciates.	N/A.	Sentence is not relevant to monetary policy.
Energy and House Prices	Oil/energy prices decrease, house prices decrease.	Oil/energy prices increase, house prices increase.	N/A.	Sentence is not relevant to monetary policy.
Foreign Nations	UK trade deficit decreases.	UK trade deficit increases.	Relating to foreign economic policy.	Sentence is not relevant to monetary policy.
BoE Expectations, Actions, Assets	Expects subpar inflation, disinflation, narrowing spreads of treasury bonds, decreases in treasury security yields, and reduction of bank reserves.	Expects high inflation, widening spreads of treasury bonds, increase in treasury security yields, increase in TIPS value, increase bank reserves.	N/A.	Sentence is not relevant to monetary policy.
Money Supply	Low money supply, slow M4 growth, increased loan demand.	High money supply, increased demand for goods, low loan demand.	N/A.	Sentence is not relevant to monetary policy.
Key Words and Phrases	Stance is “accommodative,” indicating a focus on “maximum employment” and “price stability.”	Indicative a focus on “price stability” and “sustained growth.”	Usage of “mixed,” “moderate,” “reaffirmed.”	Sentence is not relevant to monetary policy.
Labor	Productivity increases, unemployment decreases, wages increase, job openings increase.	Productivity decreases, unemployment increases, wages decrease, job openings decrease.	N/A.	Sentence is not relevant to monetary policy.

1748 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 41 and whether the committee is talking about the sentence with Certainty attest
1749 to it or not in Table 42.
1750

Table 41: Bank of England Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Data or statements that focus on anticipated future economic conditions, forecasts, or expectations, which inform potential policy actions or market movements.	“The upcoming Budget may be expansionary.”
Not Forward Looking	Data or analysis that reflects past economic performance or trends, used to understand historical conditions.	“Retail sales volumes had fallen sharply in April, in part reflecting weather-related volatility in spending.”

Table 42: Bank of England Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Information or outcomes that are highly reliable and predictable, with little room for doubt or uncertainty in economic forecasts or policy guidance.	“That would reduce lending rates for businesses with floating-rate loans linked to LIBOR.”
Uncertain	Data, forecasts, or situations that involve ambiguity or risk, where the outcome is unpredictable or subject to significant variability, affecting decision-making.	“It seemed possible that a further broad-based monetary stimulus would on its own be insufficient to transform the outlook for growth.”

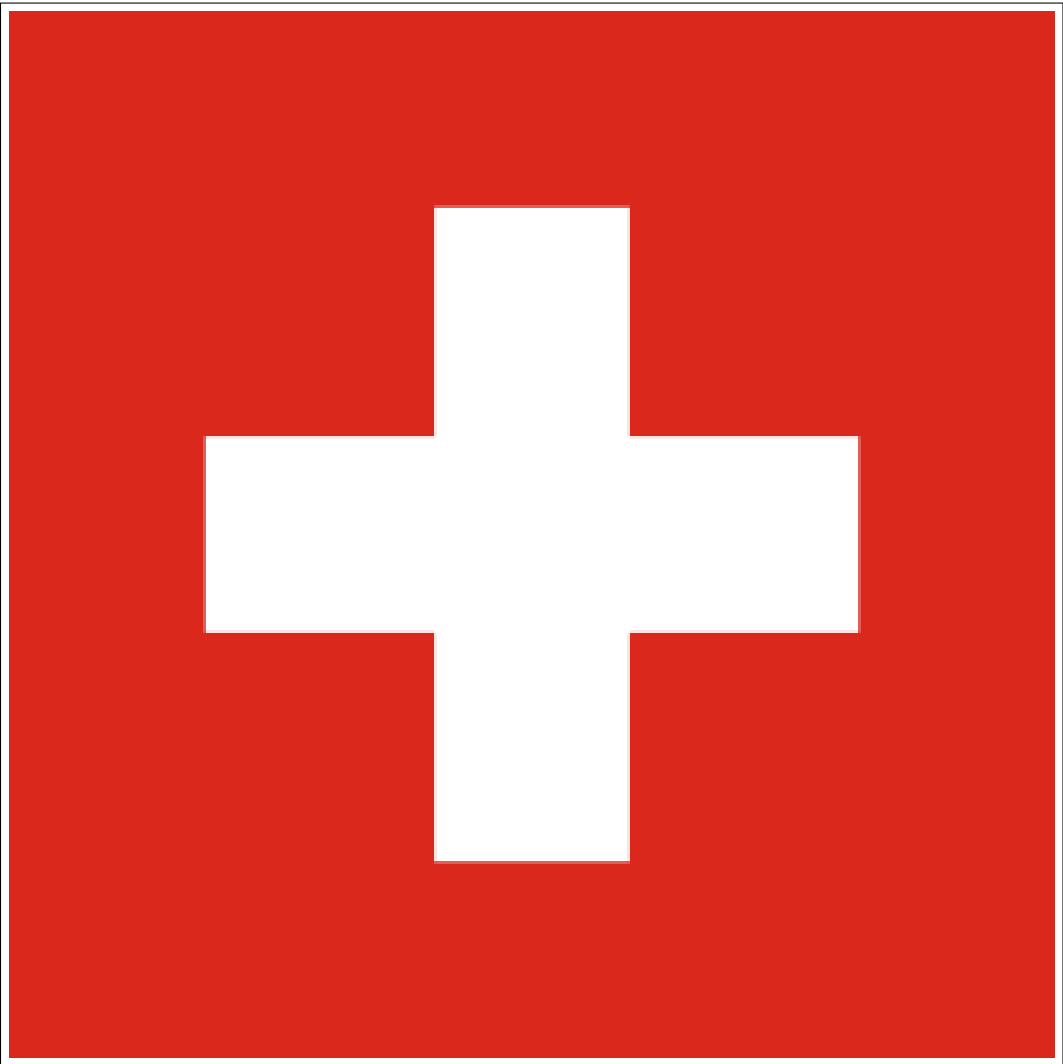
1751

Swiss National Bank

1752

Region: Switzerland

1753



1754

1755

Data Collected: 2000-2024

1756

1757

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

1758 Monetary Policy Mandate

1759 According to the Swiss National Bank's (SNB) official website, their mandate is to "ensure price
1760 stability. In so doing, it shall take due account of economic developments"

1761 Mandate Objectives:

- 1762 • **Price Stability:** Aiming for between 0-2% annual inflation rate, as measured by the National
1763 Consumer Price Index (CPI), to maintain purchasing power while considering the economy's
1764 economic situation as a whole.

1765 Structure

1766 Composition:

- 1767 • **The Governing Board:** The highest governing body in the SNB made up of the Chairperson,
1768 Vice-Chairperson, and one additional member.
- 1769 • **The Enlarged Governing Board:**
 - 1770 – **Permanent Members:** The Governing Board
 - 1771 – **Alternate Members:** Up to six alternate members (two per governing member) who
1772 are recommended by the Bank Council and chosen by the Federal Council, Switzerland's
1773 highest governing body.
- 1774 • **The Bank Council:** Made up of eleven members, six chosen by the Federal Council and five
1775 chosen by shareholders
 - 1776 – **Committees:** Comprises of the Risk Committee, Audit Committee, Compensation
1777 Committee, and Nomination Committee.
- 1778 • **Shareholders:** Elect five members of the Bank Council and vote on profit distribution.

1779 Meeting Structure:

- 1780 • **Frequency:** Quarterly meetings to determine their monetary policy and set interest rates
- 1781 • **Additional Meetings:** An annual shareholder meeting in April

1782 Manual Annotation

1783 Annotators

- 1784 • Pranav Aluru
- 1785 • Sahasra S. Chava
- 1786 • Arnav Hiray
- 1787 • Dylan Kelly

1788 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 1789 • Stance Agreement: 60.3%
- 1790 • Time Agreement: 78.0%
- 1791 • Certainty Agreement: 83.4%

1792 **Annotation Guide** Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify
1793 their annotation guide further, keeping in mind the changing global macroeconomic dynamics. Also,
1794 we have two additional categories (Forward Looking or Not Forward Looking and Certain or
1795 Uncertain) of labels other than monetary policy stance.

1796 The annotation guide for monetary policy stance is summarized in Table 43. It is built by dividing
1797 each target sentence into Nine defined categories: Interest Rates, Inflation Rate, GDP Growth
1798 Forecast, Export Performance, Consumer Demand, Employment Levels, Exchange Rate Movements,
1799 Banking Industry, and Tourism.

- 1800 • *Interest Rates:* A sentence pertaining to changes in interest rates or the Swiss Average Rate
1801 Overnight (SARON).
- 1802 • *Inflation Rate:* A sentence pertaining to the inflation rate, consumer price index, or some
1803 other inflation-related metric or references price stability.
- 1804 • *GDP Growth Forecast:* A sentence pertaining to future economic growth expectations.
- 1805 • *Export Performance:* A sentence pertaining to the export level
- 1806 • *Consumer Demand:* A sentence pertaining to consumer demand and consumer spending
1807 level
- 1808 • *Employment Levels:* A sentence pertaining to employment
- 1809 • *Exchange Rate Movements:* A sentence pertaining to the value of the Swiss Franc.
- 1810 • *Banking Industry:* A sentence pertaining to Switzerland's banking industry.
- 1811 • *Tourism:* A sentence pertaining to the tourism level in Switzerland.

1812 **Examples:**

- 1813 • "The economy has remained relatively robust despite the deterioration in the world economy,
1814 but the slowdown is likely to continue over the months to come."
1815 **Dovish:** The economic is projected to slowdown so the SNB will be inclined to promote
1816 spending.
- 1817 • "Since the indicators for the demand for labour are still at a high level, the rise in employment
1818 is likely to continue in the first part of 2008."
1819 **Hawkish:** The SNB expects employment to increase causing more money to enter into the
1820 economy.
- 1821 • "Inflationary pressure from abroad will remain weak."
1822 **Neutral:** This statement does not directly comment on the state of the Swiss economy.
- 1823 • "Through to the second quarter of 2007, this curve is a little higher than for June."
1824 **Irrelevant:** It is unclear what this statement is referring to.

Table 43: Swiss National Bank Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Interest Rates	When interest rate or SARON rate is high.	When interest rate or SARON rate is low.	When interest rate or SARON rate is stable.	Sentence is not relevant to monetary policy.
Inflation Rate	When inflation more than the target of 2%.	When inflation is below the desired level.	When is at the target of 0-2%.	Sentence is not relevant to monetary policy.
GDP Growth Forecast	When GDP is increasing.	When GDP is decreasing.	When GDP growth rate is stable.	Sentence is not relevant to monetary policy.
Export Performance	When exports are performing better than expected and earnings are rising.	When exports are declining and earnings are falling.	When exports are as expected and the rate is sustainable.	Sentence is not relevant to monetary policy.
Consumer Demand	When consumer demand is higher than expected and inflation is increasing.	When consumer demand is weak and spending is low a.	When the economy is growing at a sustainable rate in line with expectations.	Sentence is not relevant to monetary policy.
Employment Levels	When unemployment is very low.	When unemployment is high.	Unemployment is at a sustainable level and within projections.	Sentence is not relevant to monetary policy.
Exchange Rate Movements	When Franc is depreciating.	When currency is appreciating.	When currency value is stable.	Sentence is not relevant to monetary policy.
Banking Industry	When the banking industry is growing too quickly.	When the global economy is weak and financial activity is low, harming the bank-dependent Swiss economy.	When the banking industry is stable.	Sentence is not relevant to monetary policy.
Tourism	When the tourism season is very active.	When tourism levels are weak.	When tourism levels are as expected.	Sentence is not relevant to monetary policy.

1825 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 44 and whether the committee is talking about the sentence with Certainty attest
1826 to it or not in Table 45.
1827

Table 44: Swiss National Bank Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	When the sentence discusses expectations, projections, or anticipations about future economic conditions, policy actions, or resource utilization.	“Although the pace of economic recovery is likely to be moderate for a time, the Committee anticipates a gradual return to higher levels of resource utilization in a context of price stability.”
Not Forward Looking	When the sentence reflects on recent or past economic data, trends, or events to describe what has already occurred.	“Commodity prices had been mixed recently after trending down earlier.”

Table 45: Swiss National Bank Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	When the sentence reflects a clear, unanimous, or definitive decision or expectation about future actions or outcomes without ambiguity.	“To support the Committee’s decision to raise the target range for the federal funds rate, the Board of Governors of the Federal Reserve System voted unanimously to raise the interest rate paid on reserve balances to 0.”
Uncertain	When the sentence expresses concerns, potential risks, or varying opinions, often using words like “could,” “may,” or “some participants.”	“Some participants were concerned that inflation could rise as the recovery continued, and some business contacts had reported that producers expected to see an increase in pricing power over time.”

1828

Central Bank of Brazil

1829

Region: Brazil

1830



1831

1832

Data Collected: 2000-2024

1833

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

1834

1835 Monetary Policy Mandate

1836 The main responsibility of the Monetary Policy Committee (Copom) of the Banco Central do Brasil
1837 (BCB) is to set the target for the Selic rate (the policy rate).

1838 Mandate Objectives:

- 1839 • **Inflation Target:** National Monetary Council (CMN) establishes the annual inflation target
1840 in the Brazilian economy, which the BCB's aims to achieve through monetary policy
1841 decisions. As of January 2025, the inflation target is 3%, with a tolerance of $\pm 1.5\%$
- 1842 • **Sustainable Economic Growth:** By maintaining a low inflation, the economy can flourish
1843 with stable prices, thus, creating employment and enhancing the society's well-being.

1844 Structure

1845 All members of Copom are selected by the President of the Republic following their approval by the
1846 Federal State, subject to the Federal Constitution.

1847 Composition:

- 1848 • **Board of Governors:** Nine members appointed by the President and confirmed by the
1849 Federal Senate, serving 4-year terms. The Board of Governors consist of
 - 1850 – BCB Governor: Member of Copom who has the casting vote.
 - 1851 – Deputy Governors: Members of the BCB who oversee many divisions of the bank such
1852 as Economic Policy, International Affairs, etc.

1853 Meeting Structure:

- 1854 • **Frequency:** Eight regular meetings a year (every 45 days) that take place over two days
1855 (two sessions).

1856 Manual Annotation

1857 Annotators

- 1858 • Rudra Gopal
- 1859 • Joshua Zhang
- 1860 • Spencer Gosden
- 1861 • Saketh Budideti

1862 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 1863 • Stance Agreement: 43.4%
- 1864 • Time Agreement: 75.6%
- 1865 • Certainty Agreement: 78.4%

1866 **Annotation Guide** Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify
1867 their annotation guide further, keeping in mind the changing global macroeconomic dynamics. Also,
1868 we have two additional categories (Forward Looking or Not Forward Looking and Certain or
1869 Uncertain) of labels other than monetary policy stance.

1870 The annotation guide for monetary policy stance is summarized in Table 46. It is built by dividing
1871 each target sentence into twelve defined categories: Selic Rate, Monetary Policy, Inflation, Aggre-
1872 gate Demand, Food Prices, Industrial Goods Prices, Services Prices, Labor Market, GDP Growth,
1873 Exchange Rate, Trade Deficit, and the Cattle Cycle.

- 1874 • *Selic Rate:* The benchmark interest rate set by Brazil's central bank.

- 1875 • *Monetary Policy*: The process by which the central bank controls the money supply, typically
1876 through adjustments to interest rates.
- 1877 • *Inflation*: The rate at which the general level of prices for goods and services is rising,
1878 reflecting changes in purchasing power.
- 1879 • *Aggregate Demand*: The total demand for goods and services in an economy at a given
1880 overall price level and period.
- 1881 • *Food Prices*: The market prices of food products.
- 1882 • *Industrial Goods Prices*: The prices of manufactured goods produced in industrial settings.
- 1883 • *Services Prices*: The prices charged for common services.
- 1884 • *Labor Market*: The interaction between workers and employers regarding employment,
1885 wages, and job availability.
- 1886 • *GDP Growth*: The rate of increase in a economy's Gross Domestic Product, indicating
1887 economic expansion.
- 1888 • *Exchange Rate*: The value of the Brazilian Real in relation to foreign currencies.
- 1889 • *Trade Deficit*: A situation where a economy's imports exceed its exports.
- 1890 • *Cattle Cycle*: The cyclical fluctuations in beef supply and prices driven by changes in cattle
1891 production.

1892 Examples

- 1893 • “Copom unanimously decided to increase the Selic rate by 0.50 p.p. to 11.25% p.a.”
1894 **Hawkish**: Increasing the interest rate is a contractionary move aimed at curbing inflation,
1895 indicating that policymakers are concerned about an overheating economy or rising inflation
1896 pressures.
- 1897 • “Copom assessed that the scenario—marked by resilient economic activity, labor market
1898 pressures, a positive output gap, increased inflation projections, and deanchored expecta-
1899 tions—requires a more contractionary monetary policy.”
1900 **Hawkish**: The sentences describes a condition (a positive output gap and rising infla-
1901 tion expectations) that justify tightening monetary policy to prevent the economy from
1902 overheating.
- 1903 • “All members agreed that it was appropriate to reduce the Selic rate by 0.50 percentage
1904 points.”
1905 **Dovish**: A reduction in the interest rate is an expansionary action designed to stimulate eco-
1906 nomic activity. It reflects a concern that the economy might be underperforming, requiring
1907 expansionary monetary policy.
- 1908 • “The combination of a robust labor market, expansionary fiscal policy, and vigorous lending
1909 to households continues to support consumption and, consequently, aggregate demand.”
1910 **Dovish**: The sentence implies that the economy is receiving supportive stimulus from
1911 other sectors, reducing the immediate need for tighter monetary policy. Even though it
1912 recognizes strong economic activity, it suggests that current measures are sufficient without
1913 extra contractionary steps.
- 1914 • “The labor market, which surprised positively throughout 2022, has shown resilience, with a
1915 net increase in job creation and relative stability in the unemployment rate.”
1916 **Neutral**: The sentence reports economic conditions—namely, stability in the labor mar-
1917 ket—without suggesting any particular policy action. It serves as an observation of the
1918 current economic state rather than a signal for change.
- 1919 • “However, most of the committee members judged that this interpretation still seems prema-
1920 ture and needs further corroboration by data.”
1921 **Irrelevant**: This statement mainly expresses skepticism toward an interpretation and does
1922 not signal any explicit monetary policy direction or stance.

Table 46: Central Bank of Brazil Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Selic Rate	Increase in Selic rate to combat inflation.	Decrease in Selic rate to combat inflation.	Selic rate remains unchanged.	Sentence is not relevant to monetary policy.
Monetary Policy	Central bank institutes contractionary or tightening monetary policy.	Central bank institutes expansionary or loosens monetary policy.	Monetary policy remains unchanged.	Sentence is not relevant to monetary policy.
Inflation	Inflation exceeds upper limit of the tolerance band. Increasing inflation, inflationary pressures or expectations of inflation.	Inflation drops below lower limit of tolerance band. Decreasing inflation, inflationary pressures or expectations of inflation.	Inflation is within the central bank's target range, with inflation and expectations stable. COPOM's tolerance range for inflation is between 2% and 5%, with the target in the middle.	Sentence is not relevant to monetary policy.
Aggregate Demand (AD)	Increasing aggregate demand.	AD is insufficient or decreasing.	Balanced AD and Aggregate Supply (AS), stable inflation, and steady growth.	Sentence is not relevant to monetary policy.
Food Prices	Increasing or high food prices.	Decreasing or low food prices.	No change in food prices.	Sentence is not relevant to monetary policy.
Industrial Goods Prices	Increasing or high price of industrialized goods.	Decreasing or low price of industrialized goods.	No change in price of industrialized goods.	Sentence is not relevant to monetary policy.
Services Prices	Increasing or high service prices.	Decreasing or low service prices.	No change in price of services.	Sentence is not relevant to monetary policy.
Labor Market	Low or decreasing unemployment. Demand for workers exceeds available supply. Excessive wage growth. Creation of more jobs.	High or decreasing unemployment. Weak demand for labor; underemployment; stagnant wages.	Full employment without overheating, stable wage growth, and balanced labor supply and demand. Average unemployment is around 10%.	Sentence is not relevant to monetary policy.
GDP Growth	GDP grows too quickly, accelerating beyond the economy's sustainable capacity.	Weak, negative (recession), or insufficient GDP growth to sustain full employment and price stability.	Stable GDP growth without excessive fluctuations. COPOM examines the circumstances and global context; on average, growth is around 1% quarterly.	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
Exchange Rate (BRL)	When the BRL depreciates.	When the BRL appreciates.	When a foreign currency appreciates or depreciates.	Sentence is not relevant to monetary policy.
Trade Deficit	Decrease in exports or increase in imports, leading to a higher trade deficit.	Increase in exports or decrease in imports, leading to a lower trade deficit.	N/A	Sentence is not relevant to monetary policy.
Cattle Cycle	Supply of beef is low, leading to higher prices.	Supply of beef is high, leading to lower prices.	N/A	Sentence is not relevant to monetary policy.

1923 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 47 and whether the committee is talking about the sentence with Certainty attest
1924 to it or not in Table 48.
1925

Table 47: Central Bank of Brazil Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Refers to expectations of or forecasts of metrics. Key words include “expects,” “forecast,” and “anticipates.”	“Consumer price indices will come under some upward pressure from the prices of rice, milk and dairy products and clothing prices.”
Not Forward Looking	Refers to past or ongoing events or values of metrics. Key words include “previous,” “historically,” and “recorded.”	“Compared to July 2011, the rate decreased 0.9 p.p.”

Table 48: Central Bank of Brazil Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Utilizes confirmed data or definite terms when discussing a metric. Key words include “firmly,” “will,” and “decided.”	“In the month, there was no spot intervention in the domestic exchange market.”
Uncertain	Utilizes uncertain terms to describe a metric. Key words include “suggests,” “possibility,” and “risk.”	“Even though inflation is under control, there are growing uncertainties with respect to how long this trend will continue, despite high productivity gains.”

1926
1927
1928

1929
1930

Reserve Bank of India

Region: India



Data Collected: 2016-2024

1931

1932

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

1933 Monetary Policy Mandate

1934 The Reserve Bank of India (RBI) is responsible for maintaining monetary stability and fostering
1935 economic growth.

1936 Mandate Objectives:

- 1937 • **Inflation Control:** Keeping inflation within a target range of 4% (+/-2%), as measured by
1938 the Consumer Price Index (CPI).
- 1939 • **Economic Growth:** Ensuring sufficient credit flow and liquidity to support balanced
1940 economic development.
- 1941 • **Financial Stability:** Managing risks in the financial system to ensure economic resilience.

1942 Structure

1943 The RBI's governance framework underpins its role in regulating monetary and financial stability.

1944 Composition:

- 1945 • **Central Board of Directors:** Led by the RBI Governor, with Deputy Governors and
1946 government-appointed members.
- 1947 • **Monetary Policy Committee (MPC):**
 - 1948 – Six members: Three internal members from the RBI and three external experts.
 - 1949 – Responsible for setting policy rates based on inflation and growth forecasts.

1950 Meeting Structure:

- 1951 • **Frequency:** Bi-monthly reviews of economic conditions and policy efficacy.
- 1952 • **Additional Meetings:** Convened during extraordinary economic developments.

1953 Manual Annotation

1954 Annotators

- 1955 • Priyanshu Mehta
- 1956 • Aryan Garg
- 1957 • Naman Tellakula
- 1958 • Akshar Ravichandran

1959 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 1960 • Stance Agreement: 50.9%
- 1961 • Time Agreement: 69.7%
- 1962 • Certainty Agreement: 77.4%

1963 Annotation Guide

1964 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
1965 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
1966 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
1967 other than monetary policy stance.

1968 The annotation guide for monetary policy stance is summarized in Table 49. It is built by dividing
1969 each target sentence into fourteen defined categories: Inflation, Repo Rate, Reverse Repo Rate, Cash
1970 Reserve Ratio, Statutory Liquidity Ratio, GDP Growth Forecast, Monetary Policy Measures, Export
1971 Performance, Manufacturing Activity, Consumer Demand, Employment Levels, Commodity Prices,
1972 Credit Growth, Exchange Rate:

1973	• <i>Inflation</i> : A sentence pertaining to the general increase in prices for goods and services; the decrease in the purchasing power of a currency.
1974	
1975	• <i>Repo Rate</i> : A sentence pertaining to the interest rate at which the central bank lends money to commercial banks.
1976	
1977	• <i>Reverse Repo Rate</i> : A sentence pertaining to the rate at which the central bank borrows funds from commercial banks to manage liquidity.
1978	
1979	• <i>Cash Reserve Ratio (CRR)</i> : A sentence pertaining to the percentage of a bank's deposits that must be held in reserve and not lent out.
1980	
1981	• <i>Statutory Liquidity Ratio (SLR)</i> : A sentence pertaining to the minimum percentage of a bank's liabilities that must be held in the form of liquid assets.
1982	
1983	• <i>GDP Growth Forecast</i> : A sentence pertaining to the projected rate of increase in a economy's Gross Domestic Product, indicating economic expansion.
1984	
1985	• <i>Monetary Policy Measures</i> : A sentence pertaining to actions taken by the central bank to influence the money supply and interest rates.
1986	
1987	• <i>Export Performance</i> : A sentence pertaining to the export activity of a economy in foreign markets.
1988	
1989	• <i>Manufacturing Activity</i> : A sentence pertaining to the level of production and industrial output within the manufacturing sector.
1990	
1991	• <i>Consumer Demand</i> : A sentence pertaining to the overall spending by households on goods and services.
1992	
1993	• <i>Employment Levels</i> : A sentence pertaining to job availability, unemployment, and overall labor market conditions.
1994	
1995	• <i>Commodity Prices</i> : A sentence pertaining to the market prices of primary raw materials and inputs used in production.
1996	
1997	• <i>Credit Growth</i> : A sentence pertaining to the rate at which bank lending increases, reflecting prevailing financial conditions.
1998	
1999	• <i>Exchange Rate Movements</i> : A sentence pertaining to fluctuations in the value of a economy's currency relative to other currencies.
2000	

2001 Examples

2002	• “With persistently high food inflation, it would be in order to continue with the disinflationary policy stance that we have adopted.”
2003	
2004	• Hawkish : Increasing the focus on curbing high inflationary pressures indicates a tightening stance aimed at controlling rising prices.
2005	
2006	• “I am, therefore, of the view that a reduction in the policy repo rate by conventional 25 bps will be inadequate.”
2007	
2008	• Hawkish : This sentence suggests that a minor rate cut is insufficient to address economic challenges, hinting at the need for stronger tightening measures.
2009	
2010	• “There has also been an inching down in the median 3-month and 1-year ahead inflation expectations which is also comforting.”
2011	
2012	• Dovish : The observation of falling inflation expectations supports a dovish stance, where easing measures are seen as appropriate given the lower inflation outlook.
2013	
2014	• “Export dependent industries such as textiles are not doing well.”
2015	
2016	• Dovish : Highlighting weak export performance implies that easing measures might be necessary to stimulate growth in export sectors.
2017	
2018	• “Since the Indian middle income consumer is price sensitive, profits have risen.”
2019	
2020	• Neutral : This sentence provides an observation on market conditions without signaling a clear need for either tightening or easing monetary policy.
2021	
2022	• “It need not be a concern for the MPC.”
	• Irrelevant : This statement mentioned the importance of a statement and does not discuss monetary policy or take a monetary policy stance.

Table 49: Reserve Bank of India Annotation Guide

Category Term	Hawkish	Dovish	Neutral	Irrelevant
Inflation	Inflation above target or increasing; rising inflationary pressures prompting tighter policy.	Inflation below target or decreasing; low inflation encouraging easing to support growth.	Inflation at target or stable with no policy change.	Sentence is not relevant to monetary policy.
Repo Rate	Increasing the repo rate to tighten monetary policy and curb inflation.	Decreasing the repo rate to ease monetary conditions and stimulate economic activity.	Maintaining the repo rate, reflecting balanced conditions.	Sentence is not relevant to monetary policy.
Reverse Repo Rate	Increasing the reverse repo rate to reduce liquidity by encouraging banks to park funds.	Decreasing the reverse repo rate to boost liquidity by incentivizing lending.	Maintaining the reverse repo rate, indicating steady liquidity management.	Sentence is not relevant to monetary policy.
Cash Reserve Ratio (CRR)	Increasing CRR to reduce liquidity by requiring banks to hold more reserves.	Decreasing CRR to increase liquidity by allowing more lending.	Maintaining CRR, showing unchanged reserve requirements.	Sentence is not relevant to monetary policy.
Statutory Liquidity Ratio (SLR)	Increasing SLR to force banks to hold more liquid assets, reducing lending capacity.	Decreasing SLR to allow banks to hold fewer liquid assets, increasing lending capacity.	Maintaining SLR, indicating unchanged liquidity constraints.	Sentence is not relevant to monetary policy.
GDP Growth Forecast	Projecting lower GDP growth that may signal the need for tighter policy if inflation risks persist.	Projecting higher GDP growth that may lead to easing measures to support expansion.	Projecting stable GDP growth, supporting maintenance of current policy stance.	Sentence is not relevant to monetary policy.
Monetary Policy Measures	Implementing tightening measures (e.g., raising rates) to combat inflation.	Implementing accommodative measures (e.g., lowering rates) to foster growth.	No significant policy changes, suggesting a continuation of current stance.	Sentence is not relevant to monetary policy.
Exports	Strong export growth that could overheat the economy and prompt tighter policy.	Weak export performance that may require easing measures to stimulate growth.	Stable export levels with consistent foreign demand, supporting current policy.	Sentence is not relevant to monetary policy.
Manufacturing Activity	Rapid expansion that may trigger inflationary pressures, calling for tightening.	Decline in activity suggesting easing measures to revive industrial growth.	Steady, balanced manufacturing output with no policy change indicated.	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
Consumer Demand	Surging demand that might cause demand-pull inflation and necessitate tightening.	Weak consumer demand indicating low spending, warranting easing measures.	Sustainable, balanced consumer demand with no inflationary impact.	Sentence is not relevant to monetary policy.
Employment	Very low unemployment that could lead to wage inflation and overheating, prompting tightening.	High unemployment suggesting the need for expansionary policies to stimulate job creation.	Stable employment with balanced growth, indicating no policy shift.	Sentence is not relevant to monetary policy.
Commodity Prices	Rising prices that can drive cost-push inflation, necessitating tightening measures.	Falling prices reducing inflationary pressures, supporting a dovish stance.	Stable prices contributing to predictable costs, maintaining current policy stance.	Sentence is not relevant to monetary policy.
Credit Growth	Rapid credit growth that might lead to financial instability and inflation, requiring tightening.	Slow credit growth indicating tight financial conditions, suggesting easing to encourage lending.	Moderate credit growth in line with expansion, indicating no change in policy stance.	Sentence is not relevant to monetary policy.
Exchange Rate Movements	A depreciating currency increasing import costs and inflation, possibly calling for tightening.	An appreciating currency reducing import costs and inflation, allowing for easing.	Stable exchange rates that keep external pressures in check, maintaining neutrality.	Sentence is not relevant to monetary policy.

2023 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 50 and whether the committee is talking about the sentence with Certainty attest to it or not in Table 51.

2024

2025

Table 50: Reserve Bank of India Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Future-oriented: describes expectations, projections, or anticipated trends. Key words include: “Expected to,” “Projected,” “Forecast,” and “Anticipates.”	“The monetary authority expects GDP growth to accelerate next quarter, highlighting its future-oriented perspective.”
Not Forward Looking	Past-oriented: refers to historical events or trends that have already occurred. Key words include “Has been,” “Recorded,” “Previous,” and “Historically.”	“Historical data shows that inflation remained steady over the past decade, illustrating a retrospective analysis.”

Table 51: Reserve Bank of India Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Indicates definite outcomes with a clear, committed stance. Key words include “Will,” “Is set to,” “Confirmed,” and “Decided.”	“The board has decided to maintain interest rates at 2.00%.”
Uncertain	Suggests possibilities or outcomes that are not fully confirmed. Key words include: “May,” “Might,” “Could,” and “Possibility.”	“The policy rate could be lowered if consumer demand remains weaker than expected.”

2026
2027
2028

2029

2030

European Central Bank
Region: European Union



Data Collected: 2015-2024

2031

2032

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

2033 **Monetary Policy Mandate**

2034 The European Central Bank's core mandate is to maintain price stability of the Euro, the single
2035 currency of the Eurozone. It also takes responsibility for overseeing the health of member state
2036 central banks, ensuring that risk remains low across the European banking system.

2037 **Mandate Objectives:**

- 2038 • **Price Stability:** The ECB seeks to keep the inflation rate, as measured by the Harmonised
2039 Index of Consumer Prices (HICP), as close to 2% as possible.

2040 The European Union achieves its mandate by manipulating three key interest rates under its control:
2041 the Main Refinancing Operations Interest Rate, and the Deposit Facility Interest. It also engages in
2042 Open Market Operations via asset purchase programmes.

2043 **Structure**

2044 The ECB is led by the 26-member Governing Council of the European Central Bank, which makes
2045 decisions to raise or lower the three key interest rates under its control. Rate, and the Marginal
2046 Lending Facility Interest Rate. It is also responsible for decisions regarding various quantitative
2047 easing policies (asset purchasing programmes) that may be in place.

2048 **Composition:**

- 2049 • **Governing Council:** The Governing Council of the ECB consists of six members on the
2050 Executive Board, plus the 20 governors of the national central banks of Eurozone countries
2051 who follow a rotational voting system.
- 2052 • **Executive Board:** Six members of the Governing Council form the Executive Board. These
2053 members include the President and Vice President of the ECB, and are not aligned with any
2054 particular Eurozone member nation.
- 2055 • **National Central Bank Governors:** The remaining 20 members of the Governing Council
2056 consist of the governors of the national central banks of the 20 Eurozone member states.

2057 **Meeting Structure:**

- 2058 • **Frequency:** Eight scheduled meetings annually, approximately every six weeks.
- 2059 • **Additional Meetings:** Held as needed to address urgent economic developments.

2060 **Manual Annotation**

2061 **Annotators**

- 2062 • Sebastian Jaskowski
- 2063 • Nihal Nizan
- 2064 • Abhishek Pillai
- 2065 • Chaitanya Sri Yetukuri

2066 **Annotation Agreement**

2067 The agreement percentage among the pairs of annotators for different labels.

- 2068 • Stance Agreement: 52.4%
- 2069 • Time Agreement: 74.6%
- 2070 • Certainty Agreement: 82.1%

2071 **Annotation Guide**

2072 Our annotation guide is inspired by the work of Shah et al. [2023]. We modify their annotation guide
2073 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
2074 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
2075 other than monetary policy stance.

2076 The annotation guide for monetary policy stance is summarized in Table 52. It is built by dividing
2077 each target sentence into nine defined categories: Economic Status, Euro Value Change, Energy
2078 Prices, House Prices, Foreign Nations, ECB Expectations/Actions/Assets, Money Supply, Key
2079 Words/Phrases, and Labor.

- 2080 • *Economic Status*: A sentence pertaining to the general state of the Eurozone economy,
2081 incorporating factors such as inflation trends, GDP performance, and overall growth risks.
- 2082 • *Euro Value Change*: A sentence that addresses variations in the Euro’s value relative to other
2083 global currencies, whether it appreciates or depreciates.
- 2084 • *Energy Prices*: A sentence relating to changes in energy costs within the Eurozone, including
2085 fluctuations in oil and natural gas prices.
- 2086 • *House Prices*: A sentence that discusses shifts in the residential real estate market within the
2087 Eurozone, particularly changes in housing prices.
- 2088 • *Foreign Nations*: A sentence concerning trade relations between the European Union and
2089 non-Eurozone countries, highlighting changes in trade deficits or external demand for EU
2090 goods.
- 2091 • *ECB Expectations, Actions, and Assets*: A sentence describing the European Central Bank’s
2092 economic outlook and policy responses, including interest rate adjustments and modifications
2093 to asset purchasing programs.
- 2094 • *Money Supply*: A sentence discussing the availability of monetary resources in the Eurozone,
2095 reflecting trends in bank capitalization, changes in the M2 measure, and shifts in loan
2096 demand.
- 2097 • *Key Words and Phrases*: A sentence that includes specific terminology or expressions
2098 frequently used by policymakers to indicate a dovish, hawkish, or neutral stance.
- 2099 • *Labor*: A sentence addressing labor market conditions, including fluctuations in unemploy-
2100 ment rates and changes in salaries.

2101 Examples

- 2102 • “The purchasing managers’ index for manufacturing export orders had fallen into contrac-
2103 tionary territory in march owing to the war in Ukraine and lockdowns in China.”
2104 **Dovish**: Slowing economic growth in the manufacturing sector is highlighted, signifying
2105 that the central bank may want to lower interest rates to stimulate the economy.
- 2106 • “In addition, options markets had, by and large, priced out low inflation outcomes over the
2107 following five years.”
2108 **Hawkish**: Future inflation expectations are high, suggesting that higher inflation may be to
2109 come, and the ECB may want to start contracting the money supply.
- 2110 • “Current interest rates are at levels that, maintained for a sufficiently long duration, will
2111 make a substantial contribution to this goal.”
2112 **Neutral**: The ECB reiterates that there will be no interest rate change.
- 2113 • “Labour costs had also confirmed continued moderate domestic price pressures.”
2114 **Hawkish**: High labor costs can lead to inflation, which may encourage the ECB to decrease
2115 the money supply.
- 2116 • “Data on new export orders pointed to a sharp fall of global trade in the second quarter of
2117 the year.”
2118 **Dovish**: This sentence indicates a significant decline in global trade, suggesting economic
2119 slowdown that may prompt policymakers to adopt a more dovish monetary policy stance.
- 2120 • “Against this background, it was recalled that the ECB also had secondary objectives.”
2121 **Irrelevant**: This sentence primarily references the ECB’s secondary objectives, which does
2122 not directly relate to immediate monetary policy actions.

Table 52: European Central Bank Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Economic Status	HICP Inflation increases or is higher than targeted 2% rate, Inflation risk present on high side, Eurozone GDP and/or Eurozone GDP growth increases.	HICP Inflation decreases or is lower than targeted 2% rate, Eurozone GDP and/or Eurozone GDP growth decreases, Economic growth risk present on the downside.	Inflation rate remains stable around 2%, Economic growth remains unchanged.	Sentence is not relevant to monetary policy.
Euro Value Change	Euro depreciates relative to other global currencies.	Euro appreciates relative to other global currencies.	Minimal or no change in the Euro Value relative to other global currencies.	Sentence is not relevant to monetary policy.
Energy Prices	Energy prices in the Eurozone, including oil and or natural gas prices, increase.	Energy prices in the Eurozone, including oil and or natural gas prices, decrease.	N/A	Sentence is not relevant to monetary policy.
House Prices	Housing prices in the Eurozone increase.	Housing price in the Eurozone decrease.	N/A	Sentence is not relevant to monetary policy.
Foreign Nations	European Union trade deficit increases.	European Union trade deficit decreases, external (to Eurozone) demand for EU goods falls.	When foreign nation outside of the Eurozone or European Union is discussed.	Sentence is not relevant to monetary policy.
ECB Expectations, Actions, and Assets	Inflation, Economic Growth higher than expected, Increase in key ECB interest rates, Selling off of Securities purchased through Purchasing Programs).	Inflation, Economic Growth lower than expected, Decrease in key ECB interest rates, Increase in or continued exercise of purchasing programs.	Inflation, Economic Growth near expectations, Unchanged ECB interest rates (Deposit Facility, Fixed Rate Tender, Marginal Lending), no change in purchasing program policy.	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
Money Supply	Money supply is high, High Capitalization of banks, M2 decreases, Low demand for loans.	Money supply is low, Low capitalization for banks, M2 increases, Increased demand for loans. Ex. "Banks reported, on balance, an increase in firms' demand for loans or drawing of credit lines in the second quarter of 2022".	N/A	Sentence is not relevant to monetary policy.
Key Words and Phrases	Indicating a focus on "price stability" and "sustained growth."	When the stance is "accommodative," a focus on "maximum employment."	Use of phrases like "mixed," "moderate," "reaffirmed."	Sentence is not relevant to monetary policy.
Labor	When unemployment rate decreases or earnings increase.	When unemployment rate increases or earnings decrease.	Unemployment rate or growth is unchanged.	Sentence is not relevant to monetary policy.

2123 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 53 and whether the committee is talking about the sentence with Certainty attest
2124 to it or not in Table 54.
2125

Table 53: European Central Bank Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	The statement makes claims or statements about possible or certain future economic conditions.	“Looking further ahead, in the absence of new disruptions, energy costs were expected to stabilize and supply bottlenecks to ease.”
Not Forward Looking	The statement makes claims or statements about current or past economic conditions (regardless of whether these conditions are certain or possible).	“Clear signs of a recovery in consumption had emerged since may, but spending remained far below pre-lockdown levels.”

Table 54: European Central Bank Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	The speaker/writer is confident in the truth, accuracy, or occurrence of the statement.	“Annual average hicp inflation for 2014 was 0.4%.”
Uncertain	The speaker/writer expresses doubt, probability, or speculation about the information.	“Growth could be higher if inflation came down more quickly than expected, if rising real incomes meant that spending increased by more than anticipated, or if the world economy grew more strongly than expected.”

2126

Central Bank of the Russian Federation

2127

Region: Russia

2128



2129

2130

Data Collected: 2014-2024

2131

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

2132

2133 Monetary Policy Mandate

2134 The CBR is responsible for formulating monetary policy in collaboration with the Russian Federation
2135 Government to meet its primary goal of maintaining the price stability of the Ruble.

2136 Mandate Objectives:

- 2137 • **Price Stability:** main priority is to ensure price stability of the Ruble and maintain sus-
2138 tainably low inflation at an annual rate of 4%, measured by the Consumer Price Index
2139 (CPI).

2140 Structure

2141 Composition:

- 2142 • **Governor:** The CBR Governor is appointed by the President of the Russian Federation and
2143 serves a four-year term.
- 2144 • **Board of Directors:** 14 committee members who are appointed by the State Duma, the
2145 President of the Russian Federation, and the CBR Governor for a five-year period.

2146 Meeting Structure:

- 2147 • **Frequency:** The CBR holds eight meetings annually split into core and interim meetings.
- 2148 • **Core Meetings:** Four quarterly meetings where the CBR publishes its medium term macroe-
2149 conomic forecast, monetary policy report, and a mandatory press release.
- 2150 • **Interim Meetings:** Four meetings held in-between the core meetings. CBR board members
2151 meet with representatives of foreign financial institutions, analysts, and community experts
2152 community to answer questions regarding monetary policy and to receive feedback from the
2153 audience.

2154 Manual Annotation

2155 Annotators

- 2156 • Siddhartha Somani
- 2157 • Rohan Bhasin
- 2158 • Dhruv Adha
- 2159 • Andrew DiBiasio

2160 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 2161 • Stance Agreement: 49.8%
- 2162 • Time Agreement: 78.7%
- 2163 • Certainty Agreement: 89.0%

2164 Annotation Guide

2165 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
2166 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
2167 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
2168 other than monetary policy stance.

2169 The annotation guide for monetary policy stance is summarized in Table 55. It is built by dividing each
2170 target sentence into thirteen defined categories: Economic Status, Ruble Value Change, Oil/Energy
2171 Prices, House Prices, Agriculture/Food Prices, Foreign Nations & Sanctions, Military/Defense, CBR
2172 Expectations/Actions/Assets, Inflation Rates (CPI), Money Supply, Key Rate Fluctuations, Key
2173 Words/Phrases, and Labor.

- 2174 • *Economic Status*: A sentence pertaining to the state of the economy, relating to inflation
2175 levels or economic growth/decay.
- 2176 • *Ruble Value Change*: A sentence pertaining to the appreciation, depreciation, or maintenance
2177 of the Ruble.
- 2178 • *Oil/Energy Prices*: A sentence that discusses fluctuations in Russian crude oil export prices,
2179 petroleum development industry performance, natural gas production and distribution,
2180 energy commodities, or the oil/energy sector as a whole within the Russian Federation.
- 2181 • *House Prices*: A sentence pertaining to changes in prices of real estate, domestic/global
2182 housing demand, mortgage rates and rental market prices.
- 2183 • *Agriculture/Food Prices*: A sentence pertaining to the performance of the agricultural sector
2184 (crop yield, food exports) and/or fluctuations in food prices.
- 2185 • *Foreign Nations & Sanctions*: A sentence pertaining to international trade relations and/or
2186 changes regarding Western sanctions on the Russian Federation and Russian response.
- 2187 • *Military/Defense*: A sentence pertaining to defense spending and/or its effect on GDP, or
2188 about labor shortages exacerbated by wartime recruitment efforts.
- 2189 • *CBR Expectations/Actions/Assets*: A sentence that discusses changes in the Bank of Russia's
2190 key rate, bond yields, foreign exchange reserves, or any other monetary policy instrument or
2191 financial asset value.
- 2192 • *Inflation Rates (CPI)*: A sentence regarding the inflation rate as measured by the Consumer
2193 Price Index (CPI) and any inflationary/dis-inflationary pressures.
- 2194 • *Money Supply*: A sentence that overtly discusses impact to the money supply or changes in
2195 demand for the Russian Federation.
- 2196 • *Key Rate Fluctuations*: A sentence that describes decisions to manipulate the Key Rate and
2197 its corresponding fluctuations.
- 2198 • *Key Words/Phrases*: A sentence containing certain key words that would classify as hawkish,
2199 dovish, or neutral, based upon its frequency and sentiment.
- 2200 • *Labor*: A sentence that relates to changes in labor productivity of the Russian Federation.

2201 **Examples:**

- 2202 • "In February—March 2019, inflation is holding somewhat lower than the Bank of Russia's
2203 expectations."
2204 **Dovish**: Low inflation as compared to expectations indicates a raise in price level is needed
2205 to meet the target, signaling that quantitative easing should be implemented to promote price
2206 growth.
- 2207 • "Moreover, inflation expectations of households increased in July."
2208 **Hawkish**: Inflationary expectations increasing signifies a potential need to implement tighter
2209 policies to bring inflation back down.
- 2210 • "The balance of risks remains skewed towards pro-inflationary risks, especially over a short-
2211 term horizon, driven by the VAT increase and price movements in individual food products."
2212 **Hawkish**: Risks being aligned with elevated pro-inflationary risks over the short term
2213 indicates a need for immediate tightening policy to decrease pro-inflationary pressures.
- 2214 • "External inflation is stable and does not exert a noticeable influence on domestic prices."
2215 **Neutral**: Inflation is stable and therefore no tightening to decrease inflation or easing to
2216 increase inflation is needed.
- 2217 • "While assessing evolving inflation dynamics and economic developments against the
2218 forecast, the Bank of Russia admits the possibility of cutting the key rate gradually in
2219 coming Q2-Q3."
2220 **Dovish**: Cutting the key rate to adjust inflation in line with the target projections is a form
2221 of monetary easing.
- 2222 • "On 27 October 2023, the Bank of Russia Board of Directors decided to increase the key
2223 rate by 200 basis points to 15.00% per annum."
2224 **Hawkish**: Increasing the key rate is a form of quantitative tightening and indicates a stricter
2225 monetary policy stance.

- 2226 • “On 3 February 2017, the Bank of Russia Board of Directors decided to keep the key rate at
2227 10.00% p.a.”
2228 **Neutral:** Maintaining the key rate indicates a neutral policy stance as no change is needed.
2229 • “2 starting from 01.07.2016 auctions were discontinued.”
2230 **Irrelevant:** This sentence mentions that auctions were discontinued it does not provide any
2231 context related to monetary policy or a monetary policy stance.

Table 55: Central Bank of Russia Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Economic Status	When inflation increases/rises above the 4% target, unemployment decreases, or economic growth is projected high.	When inflation decreases/falls below the 4% target, unemployment increases, or economic growth is projected low.	When the unemployment rate, growth, inflation, or key rate is unchanged.	Sentence is not relevant to monetary policy.
Ruble Value Change	When the ruble depreciates - CBR may raise rates or intervene in the forex market to support the ruble.	When the ruble appreciates - CBR may intervene to slow appreciation or cut rates to reduce attractiveness of ruble assets.	N/A	Sentence is not relevant to monetary policy.
Oil/Energy Prices	When oil/energy prices increase, when their performance in the oil/energy sector, when oil development and/or exports are high.	When oil/energy prices decrease, when their performance in the oil/energy sector, when oil development and/or exports are low.	N/A	Sentence is not relevant to monetary policy.
House Prices	When housing prices, mortgage rates, housing demand, or rental market prices increase.	When housing prices, mortgage rates, housing demand, or rental market prices decrease.	N/A	Sentence is not relevant to monetary policy.
Agriculture and Food Prices	When the performance of the agricultural sector is high (e.g high exports, large crop yield) or food prices increase.	When the performance of the agricultural sector is low (e.g low exports, low crop yield) or food prices decrease.	N/A	Sentence is not relevant to monetary policy.
Foreign Nations & Sanctions	When sanctions from other countries tighten, when trading relations worsen, increase in capital inflow, or improving alliances or counter-sanctions that directly challenge Western influence.	When sanctions from other countries ease, when trading relations improve and increase capital outflow, or seeking diplomatic resolutions or negotiations to ease sanctions pressure.	Maintaining existing geopolitical relationships without significant policy shifts and when foreign trade relations remain stable.	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
Military and Defense	Positive effect on GDP from increase in defense spending.	Negative effect on GDP from decreased defense spending.	N/A	Sentence is not relevant to monetary policy.
CBR Expectations, Actions, and Assets	CBR expects higher inflation, tightening reserve requirements.	CBR expects lower inflation, easing reserve requirements.	Expecting balanced growth and ruble/price stability.	Sentence is not relevant to monetary policy.
Inflation Rates (CPI)	CPI rises, signaling inflationary pressure.	CPI (inflation rate) falls, indicating price stability or deflationary pressures.	When CPI remains stable.	Sentence is not relevant to monetary policy.
Money Supply	Tightening money supply, stricter lending policies, targeting lending risks.	Increasing money supply, easing lending conditions, supporting increased liquidity.	N/A	Sentence is not relevant to monetary policy.
Key Rate Fluctuations	Key rate hikes to control inflation and stabilize the ruble.	Key rate cuts stimulate economic activity via monetary easing.	When the key rate is maintained between meetings/terms.	Sentence is not relevant to monetary policy.
Key Words and Phrases	Price stability, target inflation, tightening policies like rate increases or raised reserve requirements.	Economic growth, accommodating, stimulating, supporting demand, consumption, export growth, or raise in GDP.	Use of phrases such as continued, steady, monitoring, and balanced.	Sentence is not relevant to monetary policy.
Labor	When productivity decreases.	When productivity increases.	N/A	Sentence is not relevant to monetary policy.

2232 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 56 and whether the committee is talking about the sentence with Certainty attest
2233 to it or not in Table 57.
2234

Table 56: Central Bank of Russia Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Semantics that indicate events or opinions formed based on events predicted to happen in the future such as “expected” or “projected.”	“The fiscal rule will set off the impact of the oil market conditions on inflation and the domestic economic environment as a whole.”
Not Forward Looking	Semantics that indicate events or opinions formed based on reliable, known data or trends.	“January 2015 saw a continued fall in real wage growth and a sharp decline in consumer expenditures which exerts a restraining influence on the prices of goods and services.”

Table 57: Central Bank of Russia Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Semantics that indicate events or opinions formed based on reliable, known data or trends.	“The fiscal rule will set off the impact of the oil market conditions on inflation and the domestic economic environment as a whole.”
Uncertain	Semantics that indicate events or opinions formed based on more volatile or unpredictable factors, like global commodity prices, using words like “potential,” “could,” “might,” or “may.”	“The economic recovery may be held back by a slower rollout of vaccination programmes, the spread of new virus strains, and the entailing toughening of restrictions, among other factors.”

Central Bank of China (Taiwan)

Region: Taiwan

The flag of the Republic of China (Taiwan) is displayed. It features a white sun with 12 rays in the upper left corner, set against a blue background. The rest of the flag is red.

Data Collected: 2017-2024**Data Collected: 2017-2024**

Important Links

[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

2242 Monetary Policy Mandate

2243 The Central Bank of China (CBC) is responsible for maintaining monetary and financial stability in
2244 Taiwan while promoting sound economic development.

2245 Mandate Objectives:

- 2246 • **Inflation Control:** Maintaining inflation around 2%, measured by the Consumer Price
2247 Index (CPI).
- 2248 • **Economic Growth:** Supporting sustainable economic growth through efficient monetary
2249 policies.
- 2250 • **Exchange Rate Stability:** Managing the New Taiwan Dollar (NTD) to ensure balanced
2251 trade and prevent financial instability.

2252 Structure

2253 Composition:

- 2254 • **Governor:** The CBC Governor oversees the implementation of monetary policies and
2255 financial operations.
- 2256 • **Board of Directors:** Comprising of 15 experts from economics, banking, and finance to set
2257 policy priorities.

2258 Meeting Structure:

- 2259 • **Frequency:** Quarterly meetings to assess economic and monetary conditions.
- 2260 • **Additional Meetings:** Convened during periods of extraordinary economic developments
2261 or financial crises.

2262 Manual Annotation

2263 Annotators

- 2264 • Nandini Shukla
- 2265 • Vatsal Dwivedi
- 2266 • Akshar Ravichandran
- 2267 • Siddhant Sukhani

2268 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 2269 • Stance Agreement: 56.9%
- 2270 • Time Agreement: 83.3%
- 2271 • Certainty Agreement: 84%

2272 Annotation Guide

2273 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
2274 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
2275 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
2276 other than monetary policy stance.

2277 The annotation guide for monetary policy stance is summarized in Table 58. It is built by dividing
2278 each target sentence into fourteen defined categories: Interest Rates, Monetary Policy Stance,
2279 Inflation Control, Economic Growth, Exchange Rate Stability, Historical Context, GDP Growth
2280 Forecasts, Inflation Expectations, Export, Manufacturing Activity, Consumer Demand, Commodity
2281 Prices, Credit Growth, and Exchange Rate Movements.

- 2282 • *Interest Rates*: A sentence pertaining to the central bank's set interest rates that influence
2283 borrowing costs and overall economic liquidity.
- 2284 • *Monetary Policy Stance*: A sentence describing the direction and approach of the central
2285 bank's policy measures, including tightening or easing actions.
- 2286 • *Inflation Control*: A sentence relating to measures aimed at keeping consumer price inflation
2287 within target ranges to maintain price stability.
- 2288 • *Economic Growth*: A sentence addressing policies or indicators that reflect the growth of
2289 Taiwan's economy.
- 2290 • *Exchange Rate Stability*: A sentence related to the actions taken to maintain a stable value
2291 of the currency.
- 2292 • *Historical Context*: A sentence providing reference to past monetary policy actions to
2293 historical situations.
- 2294 • *GDP Growth Forecasts*: A sentence discussing projected changes in the Gross Domestic
2295 Product, which is often used as an indicator of economic performance.
- 2296 • *Inflation Expectations*: A sentence discussing inflation expectations, including actions or
2297 policy that would affect inflation expectations.
- 2298 • *Export*: A sentence pertaining to the performance of a economy's exports and its implications
2299 on the economy.
- 2300 • *Manufacturing Activity*: A sentence discussing manufacturing production, including metrics
2301 such as manufacturing output.
- 2302 • *Consumer Demand*: A sentence describing the spending of households and its influence on
2303 the broader economy.
- 2304 • *Commodity Prices*: A sentence relating to the fluctuations in prices of commodities such as
2305 raw materials and energy.
- 2306 • *Credit Growth*: A sentence discussing the rate at which lending is fluctuating, indicating
2307 potential shifts in financial stability.
- 2308 • *Exchange Rate Movements*: A sentence addressing changes in the currency exchange rate.

2309 **Examples:**

- 2310 • "CBC focuses on price stability under 2%."
2311 **Hawkish**: The sentence signals that the central bank is committed to strict inflation control,
2312 implying a hawkish stance.
- 2313 • "During the Asian Financial Crisis, Taiwan protected its currency."
2314 **Hawkish**: The sentence highlights a decisive response in a crisis, implying a tight monetary
2315 approach.
- 2316 • "Forecast CPI at 1.89% supports growth."
2317 **Dovish**: The sentence suggests a relatively low CPI, indicating that there is leeway for
2318 expansionary monetary policy.
- 2319 • "Weak orders for electronics and machinery may warrant supportive monetary steps."
2320 **Dovish**: The sentence indicates that easing measures could be necessary to bolster export
2321 demand.
- 2322 • "CBC maintained rates at 2.00% to ensure stability."
2323 **Neutral**: The sentence provides a factual statement about unchanged interest rates without
2324 directional bias.
- 2325 • "The bank successively met with a total of 34 domestic banks between august 12 and august
2326 21."
2327 **Irrelevant**: The sentence describes a series of routine meetings with domestic banks and
2328 does not provide any indication of monetary stance.

Table 58: Central Bank of China Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Interest Rates	High rates to control inflation and reduce liquidity.	Lower rates to encourage borrowing and spending.	Rates remain unchanged.	Sentence is not relevant to monetary policy.
Monetary Policy Stance	Tightening measures such as reduced asset purchases or clear signals of future rate hikes.	Accommodative measures, including rate cuts, quantitative easing, or liquidity injections.	No significant change in policy stance or rates, indicating a wait-and-see approach.	Sentence is not relevant to monetary policy.
Inflation Control	Strong measures to keep CPI well below 2%.	Tolerance for higher inflation to support economic growth.	Inflation projections stable or within comfortable bounds.	Sentence is not relevant to monetary policy.
Economic Growth	Policies to prevent overheating.	Measures to stimulate growth (e.g., rate cuts).	Balanced approach to growth.	Sentence is not relevant to monetary policy.
Exchange Rate Stability	Aggressive forex interventions to stabilize or strengthen the currency.	Allowing slight depreciation to boost exports; looser approach to FX.	Stable exchange rates are maintained.	Sentence is not relevant to monetary policy.
Historical Context	Tightening responses in past crises.	Accommodative responses in past downturns.	Mentions of periods without major policy shifts.	Sentence is not relevant to monetary policy.
GDP Growth Forecasts	Focus on curbing excessive expansion to mitigate inflation.	Policies that drive higher growth through domestic demand or exports.	Stable projections indicating moderate, sustainable growth.	Sentence is not relevant to monetary policy.
Inflation Expectations	Very strict control near 2% or lower to avoid price instability.	Willingness to let inflation rise slightly if beneficial for growth.	Projections remain moderate and steady.	Sentence is not relevant to monetary policy.
Export	Robust export growth, risking overheating and trade imbalances.	Declining export demand, suggesting easing measures to support exporters.	Stable export levels with sustainable growth patterns.	Sentence is not relevant to monetary policy.
Manufacturing Activity	Rapid output potentially causing input shortages or inflation pressures.	Decline in factory output due to weak global demand or disruptions.	Balanced output growth indicating stable demand in key markets.	Sentence is not relevant to monetary policy.
Consumer Demand	Rising domestic consumption pushing inflation beyond targets.	Weak household spending needing policy stimulus.	Stable consumer demand, supporting GDP without driving up inflation.	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
Commodity Prices	Rising global commodity costs (energy, raw materials) pressuring margins.	Falling import prices easing inflation and supporting growth.	Stable commodity prices reducing uncertainty in production and exports.	Sentence is not relevant to monetary policy.
Credit Growth	Rapid credit expansion increasing leverage and financial instability risks.	Slowing credit growth indicating tight conditions, possibly needing easing.	Moderate credit growth aligned with economic needs (e.g., SMEs).	Sentence is not relevant to monetary policy.
Exchange Rate Movements	NTD appreciation, harming exports and requiring tighter interventions.	NTD depreciation, helping exports but risking imported inflation.	Stable rates ensuring balanced trade flows with minimal inflation pressure.	Sentence is not relevant to monetary policy.

2329 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 59 and whether the committee is talking about the sentence with Certainty attest
2330 to it or not in Table 60.
2331

Table 59: Central Bank of China (Taiwan) Annotation Guide to identify whether a sentence has forward-looking information or not.

Category	Definition	Sentence Example
Forward Looking	Future-oriented: describes expectations, projections, or anticipated trends. Key words include “expected to,” “projected,” “forecast,” and “anticipates.”	“The central bank projects GDP growth of 3% next year.”
Not Forward Looking	Past-oriented: refers to historical events or trends that have already occurred. Key words include “has been,” “recorded,” “previous,” and “historically.”	“Historically, inflation has been below 2%, reflecting stable price levels in past years.”

Table 60: Central Bank of China Annotation Guide to identify whether a sentence echoes certainty or not.

Category	Definition	Sentence Example
Certain	Indicates definite outcomes with a clear, committed stance. Key words include “will,” “is set to,” “confirmed,” and “decided.”	“The central bank will raise the policy rate next month.”
Uncertain	Suggests possibilities or outcomes that are not fully confirmed. Key words include “may,” “might,” “could,” and “possibility.”	“The monetary authority may lower rates if domestic demand softens further.”

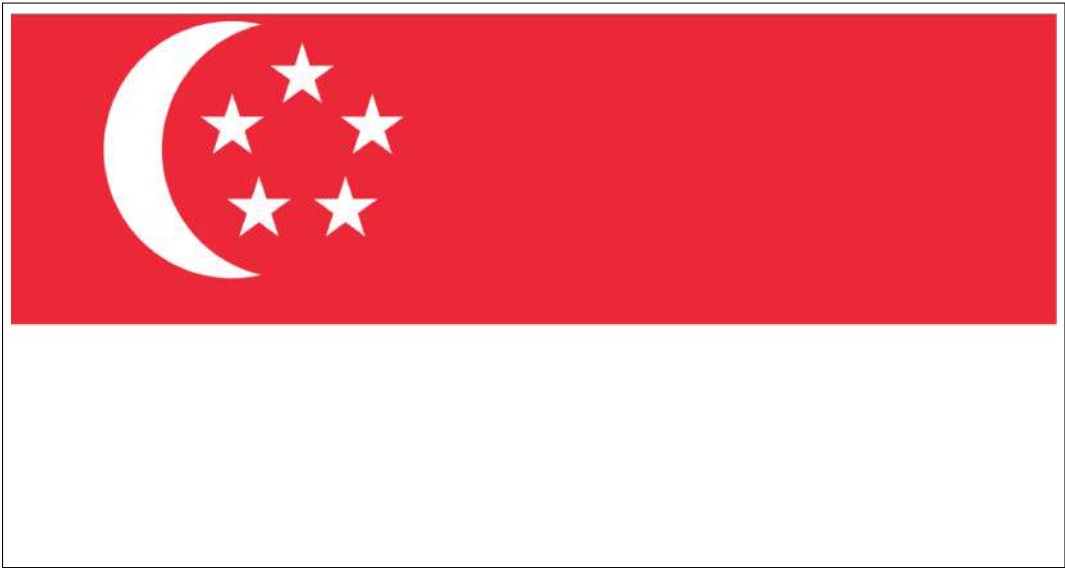
2332

Monetary Authority of Singapore

2333

Region: Singapore

2334



2335

2336

Data Collected: 2013-2024

2337

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

2338

2339 Monetary Policy Mandate

2340 The Monetary Authority of Singapore (MAS) is responsible for maintaining monetary and financial
2341 stability in Singapore while fostering economic growth.

2342 Mandate Objectives:

- 2343 • **Exchange Rate Policy:** Managing the Singapore dollar (SGD) for price stability and
2344 economic growth.
- 2345 • **Inflation Control:** Controlling inflationary pressures using exchange rate tools.
- 2346 • **Financial Stability:** Ensuring the resilience of Singapore’s financial system.

2347 Structure

2348 Composition:

- 2349 • **Managing Director:** Oversees policy implementation and financial operations.
- 2350 • **Board of Directors:** Composed of economic and financial experts setting strategic policy
2351 directions.

2352 Meeting Structure:

- 2353 • **Frequency:** Regular policy reviews to assess macroeconomic conditions.
- 2354 • **Additional Meetings:** Held during significant economic events.

2355 Manual Annotation

2356 Annotators

- 2357 • Marc-Alain Adjahi
- 2358 • Jessica Ly
- 2359 • Vivek Atmuri
- 2360 • Giovanni Hsu

2361 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 2362 • Stance Agreement: 47.5%
- 2363 • Time Agreement: 91.9%
- 2364 • Certainty Agreement: 79%

2365 Annotation Guide

2366 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
2367 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
2368 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
2369 other than monetary policy stance.

2370 The annotation guide for monetary policy stance is summarized in Table 61. It is built by dividing
2371 each target sentence into six defined categories: Exchange rate policy, Inflation control, Employment,
2372 Trade, Foreign labor and Talent attraction, and Global Supply Chain Hub.

- 2373 • *Exchange Rate Policy:* A sentence pertaining to changes such as the appreciation or depreci-
2374 ation of the value of the Singapore dollar (SGD).
- 2375 • *Inflation Control:* A sentence pertaining to adjusting the inflationary pressure by weakening
2376 or strengthening the SGD.
- 2377 • *Employment:* A sentence pertaining to changes in the employment sectors.

- 2378 • *Trade*: A sentence pertaining to trade relations between Singapore and a foreign economy.
2379 If not discussing Singapore, we label irrelevant.
- 2380 • *Foreign Labor and Talent Attraction*: A sentence that discusses changes in immigration
2381 policies to increase or limit foreign talent.
- 2382 • *Global Supply Chain Hub*: A sentence that discusses Singapore's involvement in the global
2383 supply chain.

2384 **Examples:**

- 2385 • "To combat rising inflation, the central bank may need to steepen the exchange rate band."
2386 **Hawkish**: Increasing the exchange rate band indicates a tighter monetary policy stance to
2387 maintain price stability.
- 2388 • "A flattening of the band would provide greater flexibility during economic downturns."
2389 **Dovish**: Flattening the exchange rate band indicates a more accommodative stance to
2390 support the economy.
- 2391 • "The current monetary policy remains effective without any proposed changes."
2392 **Hawkish**: There are no changes made to the monetary policy and remains neutral.
- 2393 • "External inflationary pressures are likely to stay muted for the rest of 2003."
2394 **Dovish**: The statement suggests that inflationary pressures are not expected to worsen,
2395 thereby creating room for expansionary monetary policy.
- 2396 • "MAS will be shifting the schedule of its semi-annual cycle from January/July to
2397 April/October."
2398 **Irrelevant**: This sentence relates to an operational or scheduling change and does not
2399 mention monetary policy decisions or stances.

Table 61: Monetary Authority of Singapore Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Exchange Rate Policy	In favor of or advocating for the appreciation of the Singapore dollar (SGD).	In favor of or advocating for the depreciation or weakening of the Singapore dollar (SGD).	Identifying economic trends but not suggestive of action to the currency.	Sentence is not relevant to monetary policy.
Inflation Control	Favoring a stronger SGD to reduce inflationary pressures.	Tending towards weakening or maintaining the SGD due to the reduction of inflationary pressures.	Discusses inflation as a whole and mentions statistics but don't take a stance on it.	Sentence is not relevant to monetary policy.
Employment	Lack of employment due to the strengthening of the SGD due to other concerns.	Favoring a weaker SGD to boost export-driven employment sectors.	Discusses employment metrics without suggesting action either way.	Sentence is not relevant to monetary policy.
Trade	Favoring stricter trade policies and protective measures to guard domestically produced items against international competition.	Favoring open markets and more relaxed policies to boost exports and secure Singapore's role as a trade hub.	Discusses previous trade strategies or options for future opportunities.	Sentence is not relevant to monetary policy.
Foreign Labor Policy	Sentences discuss the need for stricter immigration policies in order to keep local jobs and limit foreign labor.	Suggesting the allowing of an increase in foreign talent in the local markets to foster innovation and economic growth.	Discusses the maintaining of immigration policies and SGD or mentioning local vs. foreign labor forces.	Sentence is not relevant to monetary policy.
Global Supply Chain	Suggesting the mitigation of global supply chain disruptions by focusing on more local sourcing.	Sentences implies the increase in involvement in global supply chain in order to enhance Singapore's position as a hub and to invite investment.	Maintaining Singapore's policy and involvement in the global supply chain due to the lack of concerns of disruptions or happiness with Singapore's place as the hub.	Sentence is not relevant to monetary policy.

2400 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 62 and whether the committee is talking about the sentence with Certainty attest
2401 to it or not in Table 63.
2402

Table 62: Monetary Authority of Singapore Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	The sentence is written in future tense or is suggestive of an action MAS will take. It could also mention a solution to a problem in the past or mention a modification to previous policies.	“To combat rising inflation, the central bank will steepen the exchange rate band.”
Not Forward Looking	The sentence states facts about a previous event or mentions action that is currently being taken. It could also discuss the pros and cons of a previous policy or reminisce on what could have been done differently.	“Previously, the flattening of the band provided greater flexibility during economic downturns.”

Table 63: Monetary Authority of Singapore Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	The sentence mentions an event that will happen, an event that is very likely to happen, or mentions an event that happened in the past.	“A flattening of the band will very likely provide greater flexibility during economic downturns.”
Uncertain	The sentence mentions an event that should happen, an event that will likely happen, an event that is unsure to happen, or an event that happened in the past, but is speculative of the reason.	“A flattening of the band should provide greater flexibility during economic downturns.”

2403

Bank of Korea

2404

Region: South Korea

2405



2406

2407

Data Collected: 2010-2024

2408

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

2409

2410 Monetary Policy Mandate

2411 The Bank of Korea (BoK) serves as South Korea's central bank, responsible for formulating and
2412 implementing monetary policy, maintaining financial stability, and issuing the national currency, the
2413 South Korean won (KRW). In addition to its policy making functions, the BoK provides banking
2414 services to the government and financial institutions, oversees payment systems, and manages the
2415 economy's foreign exchange reserves. It also conducts research on macroeconomic and financial
2416 issues to support policy decisions.

2417 **Mandate Objectives:** The functions and responsibilities of the BoK are established by various
2418 legislative frameworks, primarily the Bank of Korea Act, which defines its authority, powers, and
2419 obligations.

- 2420 • **Price Stability:** Maintaining price stability as the foundation of sustainable economic
2421 growth.
- 2422 • **Financial Stability:** Ensuring the soundness of the financial system to prevent systemic
2423 risks.
- 2424 • **Economic Growth Support:** Contributing to balanced economic development and employ-
2425 ment stability.
- 2426 • **Foreign Exchange Management:** Managing foreign reserves and exchange rate policies to
2427 maintain external stability.

2428 Structure

2429 The Bank of Korea (BoK) Monetary Policy Board is responsible for formulating and implementing
2430 monetary policy. Before each meeting, BoK staff prepare detailed reports on domestic and interna-
2431 tional economic conditions, financial market developments, and policy recommendations. Senior
2432 officials present these analyses during meetings to inform decision-making.

2433 Composition:

- 2434 • **Executive Members:** The Monetary Policy Board consists of the following key officials:
 - 2435 – The Governor (Chair): serves a four-year term and may be reappointed for a single
2436 consecutive term
 - 2437 – The Senior Deputy Governor: serves a three-year term and may be reappointed for a
2438 single consecutive term
 - 2439 – The Deputy Governors: serves a three-year term and may be reappointed for a single
2440 consecutive term
- 2441 • **Non-Executive Members:** Additional members are appointed by the President of South
2442 Korea upon the recommendation of the government and financial sector authorities.

2443 Meeting Structure:

- 2444 • **Frequency:** Eight times a year, following the release of key economic indicators such as
2445 inflation reports, GDP growth figures, and financial stability assessments. Additionally, the
2446 Board convenes regularly on the second and fourth Thursdays of each month to discuss
2447 various economic and financial issues.
- 2448 • **Additional Meetings:** Held if deemed necessary by the Chairman or upon request by at
2449 least two members.

2450 Manual Annotation

2451 Annotators

- 2452 • Soungmin (Min) Lee
- 2453 • Sarah Chae

2454 • Eric Kim
 2455 • Woojin Kang

2456 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

2457 • Stance Agreement: 48.0%
 2458 • Time Agreement: 72.9%
 2459 • Certainty Agreement: 65.6%

2460 **Annotation Guide**

2461 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
 2462 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
 2463 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
 2464 other than monetary policy stance.

2465 The annotation guide for monetary policy stance is summarized in Table 64. It is built by dividing
 2466 each target sentence into 14 defined categories: : Economic Status, Keywords/Phrases, KRW Value
 2467 Change, Housing Debt/Loan, House Prices, Foreign Nations, BoK Expectations/Actions/Assets,
 2468 Money Supply, Labor, Interest rate gap between South Korea and United States, US Dollar Value,
 2469 Foreign Exchange (FX) Market, COVID-19, and Energy Prices.

2470 • *Economic Status*: A sentence pertaining to the state of the economy, relating to unemploy-
 2471 ment and inflation.

2472 • *Keywords/Phrases*: A sentence that contains key word or phrase that would classify it
 2473 squarely into one of the three label classes, based upon its frequent usage and meaning
 2474 among particular label classes.

2475 • *KRW Value Change*: A sentence pertaining to changes such as appreciation or depreciation
 2476 of value of the South Korea Won on the Foreign Exchange Market.

2477 • *Housing Debt/Loan*: A sentence that relates to changes in debt in households.

2478 • *House Prices*: A sentence pertaining to changes in prices of real estate.

2479 • *Foreign Nations*: A sentence pertaining to trade relations between South Korea and a foreign
 2480 economy. If not discussing South Korea we label neutral.

2481 • *BoK Expectations/Actions/Assets*: A sentence that discusses changes in BoK yields, bond
 2482 value, reserves, or any other financial asset value.

2483 • *Money Supply*: A sentence that overtly discusses impact to the money supply or changes in
 2484 demand.

2485 • *Labor*: A sentence that relates to changes in labor productivity.

2486 • *Interest rate gap between South Korea and United States*: A sentence pertaining to changes
 2487 such as increase or decrease of the interest rate gap between the United States and South
 2488 Korea.

2489 • *US Dollar Value*: A sentence about fluctuation of the dollar’s value which can influence
 2490 monetary policy in a dovish or hawkish direction.

2491 • *FX Market*: A sentence about stability’s changes in the foreign exchange market.

2492 • *COVID-19*: A sentence indicating whether the government is tightening or easing COVID
 2493 policy.

2494 • *Energy Prices*: A sentence pertaining to changes in prices of energy commodities or the
 2495 energy sector as a whole.

2496 **Examples:**

2497 • “Concerning prices, the member presented the view that a depreciation of the won could
 2498 further fuel inflationary pressures.”
 2499 **Hawkish:** This statement warns that a depreciation of the won may lead to increased
 2500 inflation, suggesting the need for a tighter monetary policy.

- 2501 • “Some of the members emphasized that it would be necessary to prepare against growing
2502 instability of the external sector due to the worsening global financial conditions including
2503 the rising won-dollar exchange rate, the narrowing current account surplus, and a possible
2504 foreign capital outflow owing to the widening gap between domestic and overseas interest
2505 rates.”
2506 **Hawkish:** This sentence reflects a hawkish stance by highlighting external financial instabil-
2507 ity and recommending proactive measures to counter potential inflationary pressures.
- 2508 • “Finally, in terms of financial stability, the member assessed that growth in household lending
2509 and housing prices had slowed since the second half of last year, but continued caution about
2510 the risk of a buildup of financial imbalances was required.”
2511 **Dovish:** This statement is dovish as it underscores caution and a preference for maintaining
2512 accommodative measures despite slowing growth in household lending and housing prices.
- 2513 • “In light of the narrowing interest rate gap between South Korea and the United States,
2514 some members suggested that the Bank of Korea may consider keeping interest rates low to
2515 support domestic economic recovery and encourage borrowing.”
2516 **Dovish:** This sentence advocates for lower interest rates to stimulate economic recovery,
2517 which is an expansionary monetary policy stance.
- 2518 • “The member, however, pointed out that the price path had not been showing a commensurate
2519 change.”
2520 **Neutral:** This sentence notes that the price path has not changed significantly without
2521 implying any particular monetary policy stance.
- 2522 • “The member thus emphasized the need to pay attention to this issue.”
2523 **Irrelevant:** This sentence is irrelevant because it does not provide enough context related to
2524 economic policy or market conditions relevant to the Bank of Korea.

Table 64: Bank of Korea Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Economic Status	When inflation decreases, when unemployment increases, when economic growth is projected as low.	When inflation increases, when unemployment decreases when economic growth is projected high when economic output is higher than potential supply/actual output when economic slack falls.	When unemployment rate or growth is unchanged, maintained, or sustained.	Sentence is not relevant to monetary policy.
KRW Value Change	When KRW value appreciates.	When KRW value depreciates.	N/A	Sentence is not relevant to monetary policy.
US Dollar Value	When USD is moderate or weakening.	When USD becomes stronger, as this indicates potential depreciation of KRW.	N/A	Sentence is not relevant to monetary policy.
Interest rate gap between South Korea and United States	When US interest rate is aligned with South Korea's, US not interfering with South Korea's interest rates.	When US interest rate is higher than South Korea's, potentially pressuring South Korea to raise interest rates.	N/A	Sentence is not relevant to monetary policy.
FX market	Stable FX market.	Unstable FX market.	N/A	Sentence is not relevant to monetary policy.
Housing Debt/Loan	When public or private debt level decreases, signaling more flexibility for borrowing and fiscal stimulus.	When public or private debt level decreases, signaling more flexibility for borrowing and fiscal stimulus.	N/A	Sentence is not relevant to monetary policy.
Housing Prices	When house prices decrease or expected to decrease.	When house prices increase or expected to increase.	N/A	Sentence is not relevant to monetary policy.
Foreign Nations	When the Korean trade deficit decreases.	When the Korean trade deficit increases.	When relating to a foreign nation's economic or trade policy.	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
BoK Expectations/Actions/Assets	BoK expect supbar inflation, BoK expecting disinflation, narrowing spreads of treasury bonds, decreases in treasury security yields, and reduction of bank reserves.	When BoK signals the need for tightening, raising interest rates, or increasing reserves to curb inflation.	N/A	Sentence is not relevant to monetary policy.
Keywords/Phrases	When the stance is "accommodative", indicating a focus on "maximum employment" and "price stability."	Indicating a focus on "price stability" and "sustained growth."	Use of phrases "mixed", "moderate", "reaffirmed."	Sentence is not relevant to monetary policy.
Money Supply	Money supply is low, M2 increases, increased demand for loans.	Money supply is high, increased demand for goods, low demand for loans.	N/A	Sentence is not relevant to monetary policy.
Labor	When productivity increases, solid labor market.	When productivity decreases, unstable labor market.	N/A	Sentence is not relevant to monetary policy.
COVID-19	Easing COVID-19 regulation: boost tourism industry and increased productivity.	Increased COVID-19 regulation: detrimental for tourism industry and decreased productivity.	N/A	Sentence is not relevant to monetary policy.
Energy	When oil/energy prices decrease.	When oil/energy prices increase.	N/A	Sentence is not relevant to monetary policy.

2525 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 65 and whether the committee is talking about the sentence with Certainty attest
2526 to it or not in Table 66.
2527

Table 65: Bank of Korea Annotation Guide to identify whether a sentence has forward-looking information or not.


Label	Description	Example
Forward Looking	When sentence contains words such as “expected,” “anticipated,” “projected,” “will,” or “forecasted.”	“Budget 2025 measures will provide additional support to growth.”
Not Forward Looking	When sentence contains phrase like “had been,” “previously,” or “last year.”	“Domestic headline inflation moderated in September to 1.8.”

Table 66: Bank of Korea Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	When sentence contains phrase like “will,” “must,” “certainly,” or “definitely.”	“The monetary policy statement will be released at 6 p.m. on the same day as the MPC meeting.”
Uncertain	When sentence contains phrase like “might,” “could,” “may,” “possibly,” “likely,” or “uncertain.”	“These developments may increase the risk to the outlook for inflation.”

Reserve Bank of Australia

Region: Australia

The image shows the national flag of Australia. It features a blue field with a large white seven-pointed star in the lower-left canton. In the upper-left canton is a white border enclosing a red field with a white Union Jack. In the fly, there are five white stars: one large seven-pointed star and four smaller five-pointed stars arranged in an arc.**Data Collected: 2006-2024**

Important Links

[Central Bank Website](#)

Central Bank Website

Stance Label Model

Time Label Model

Certain Label Model

2535 **Monetary Policy Mandate**

2536 The Reserve Bank of Australia (RBA) serves as Australia's central bank, responsible for conducting
2537 monetary policy, maintaining financial stability, and issuing the national currency. Beyond its
2538 policy-making role, the RBA provides select banking and registry services to Australian government
2539 agencies, overseas central banks, and official institutions. Additionally, it manages Australia's gold
2540 and foreign exchange reserves.

2541 The functions and responsibilities of the RBA are defined by various legislative frameworks, primarily
2542 the *Reserve Bank Act 1959*, which establishes its statutory authority, powers, and obligations.

2543 **Mandate Objectives** The RBA's primary objectives, as outlined in the *Reserve Bank Act 1959*,
2544 include:

- 2545 • **Currency Stability:** Ensuring the stability of the Australian dollar.
- 2546 • **Full Employment:** Promoting full employment within Australia.
- 2547 • **Economic Prosperity:** Supporting the economic welfare and prosperity of the Australian
2548 people.

2549 **Structure**

2550 The Reserve Bank Board is responsible for formulating monetary policy. Before each meeting, RBA
2551 staff prepare comprehensive reports on domestic and international economic conditions, financial
2552 markets, and policy recommendations. Senior staff attend and present during these meetings.

2553 • **Composition**

2554 – **Executive Members:** The Board comprises three *ex officio* members:

- 2555 * The Governor (Chair)
- 2556 * The Deputy Governor (Deputy Chair)
- 2557 * The Secretary to the Treasury

2558 The Governor and Deputy Governor serve terms of up to seven years and may be
2559 reappointed.

2560 – **Non-Executive Members:** Six non-executive members are appointed by the Treasurer
2561 for terms of up to five years.

2562 • **Meeting Structure:** The Board convenes eight times per year, following the release of key
2563 economic indicators, such as inflation and economic activity reports. Meetings take place at
2564 the RBA's Head Office in Sydney.

2565 • **Quorum and Decision-Making:** A quorum requires five members, with the Governor or,
2566 in their absence, the Deputy Governor presiding. Decisions are made by majority vote, with
2567 the Chair holding a casting vote if necessary. Minutes of monetary policy meetings are
2568 published two weeks after each meeting. The Governor and Deputy Governor do not partake
2569 in discussions concerning their terms of employment.

2570 • **Meeting Schedule:** Meetings commence on a Monday afternoon and conclude the following
2571 day. The outcome is announced at 2:30 PM on the second day. Following each meeting, the
2572 Governor holds a media conference to explain the policy decision.

2573 **Meeting Structure:** The Board convenes eight times per year, following the release of key economic
2574 indicators, such as inflation and economic activity reports. Meetings take place at the RBA's Head
2575 Office in Sydney.

2576 **Quorum and Decision-Making:** A quorum requires five members, with the Governor or, in their
2577 absence, the Deputy Governor presiding. Decisions are made by majority vote, with the Chair holding
2578 a casting vote if necessary. Minutes of monetary policy meetings are published two weeks after each
2579 meeting. The Governor and Deputy Governor do not partake in discussions concerning their terms of
2580 employment.

2581 **Meeting Schedule:** Meetings commence on a Monday afternoon and conclude the following day.
2582 The outcome is announced at 2:30 PM on the second day. Following each meeting, the Governor
2583 holds a media conference to explain the policy decision.

2584 Manual Annotation

2585 Annotators

- 2586 • Eric Van Ness
- 2587 • Ankit Agrawal
- 2588 • Lana Duke
- 2589 • Nithil Balaji

2590 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 2591 • Stance Agreement: 63.3%
- 2592 • Time Agreement: 88.8%
- 2593 • Certainty Agreement: 83.2%

2594 Annotation Guide

2595 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
2596 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
2597 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
2598 other than monetary policy stance.

2599 The annotation guide for monetary policy stance is summarized in Table 67. It is built by dividing
2600 each target sentence into eleven defined categories: Economic Status, Government Spending, Key
2601 Words and Phrases, Dollar Value Change, Energy, Commodity, House Prices, RBA Expectations,
2602 Actions, Assets, Foreign Nations, Money Supply, Labor, Consumer Sentiment, Wages.

2603 Definitions:

- 2604 • *Economic Status*: A sentence discussing the current state of inflation, unemployment, or
2605 overall economic growth, and how these indicators influence monetary policy decisions.
- 2606 • *Government Spending*: A sentence pertaining to levels of government spending, such as
2607 enacting expansionary or contractionary fiscal policy.
- 2608 • *Key Words and Phrases*: A sentence that contains key word or phrase that would classify
2609 squarely into one of the three label classes, based upon its frequent usage and meaning
2610 among particular label classes.
- 2611 • *Dollar Value Change*: A sentence pertaining to changes such as appreciation or depreciation
2612 of value of the Australian Dollar on the Foreign Exchange Market.
- 2613 • *Energy/Commodity/House Prices*: A sentence pertaining to Energy via changes in prices of
2614 energy commodities or the energy sector as a whole, Commodities via precious metals or
2615 agricultural goods, or House Prices via single-family homes or the real estate sector as a
2616 whole.
- 2617 • *RBA Expectations/Actions/Assets*: A sentence that discusses changes in the cash rate, bond
2618 value, reserves, or any other financial asset value.
- 2619 • *Foreign Nations*: A sentence pertaining to trade relations between Australia and a foreign
2620 economy. If not discussing Australia we label neutral.
- 2621 • *Money Supply*: A sentence that overtly discusses impact to the money supply or changes in
2622 demand.
- 2623 • *Labor*: A sentence that relates to changes in labor productivity.
- 2624 • *Consumer Sentiment*: A sentence reflecting the general confidence of households regarding
2625 future economic conditions.
- 2626 • *Wages*: A sentence highlighting wage growth or decline, which impacts inflation and
2627 consumer spending

2628 Examples:

- 2629 • “The AONIA remained below target despite a recent policy change.”
- 2630 **Dovish:** Indicates that lower-than-target overnight lending rates suggest an easing stance.
- 2631 • “The RBA’s stance has been accommodative toward recent inflationary pressures.”
- 2632 **Dovish:** Suggests that the central bank is prioritizing economic support over tightening
- 2633 measures.
- 2634 • “Inflation has proved to be persistent and a challenge, staying above the inflation target even
- 2635 with recent efforts.”
- 2636 **Hawkish:** Implies that ongoing high inflation may necessitate a tightening of monetary
- 2637 policy.
- 2638 • “Import trade restrictions for China, one of Australia’s key trading partners could lead to
- 2639 supply shortages and higher prices, consequentially increasing inflationary pressures.”
- 2640 **Hawkish:** Highlights how external trade barriers can drive inflation upward, justifying
- 2641 stricter policy.
- 2642 • “After consideration, the cash rate target will remain unchanged as of October 2024.”
- 2643 **Neutral:** Reflects a decision to maintain the current policy without alteration.
- 2644 • “Members recognized that the calibration of this guidance was not precise or straightfor-
- 2645 ward.”
- 2646 **Irrelevant:** Discusses the clarity and formulation of policy guidance rather than actual
- 2647 economic indicators or monetary actions.

Table 67: Reserve Bank of Australia Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Economic Status	When inflation decreases, when economic growth is predicted to be low, when cash rate (the goal overnight lending rate) target will be lowered, when inflation is below the inflation target, when the AONIA (the overnight lending rate) is below the cash rate target.	When inflation increases, when economic growth is predicted to be high, when cash rate target will be raised, when inflation is above the inflation target, when the AONIA (aka cash rate) is above the cash rate target.	When inflation remains unchanged, when the cash rate target remains unchanged, when inflation is at the inflation target, when the cash rate is at the cash rate target.	Sentence is not relevant to monetary policy.
Government Spending	When government spending decreases.	When government spending increases.	N/A	Sentence is not relevant to monetary policy.
Key Words and Phrases	When the stance is “accommodative,” indicating a focus on “maximum employment” and “price stability.”	Indicating a focus on “price stability” and “sustained growth.”	Use of phrases “mixed,” “moderate,” “reaffirmed.”	Sentence is not relevant to monetary policy.
Dollar Value Change	When the Australian Dollar appreciates.	When the Australian Dollar depreciates.	N/A	Sentence is not relevant to monetary policy.
Energy, Commodity, House Prices	When energy, commodities, or home prices decrease.	When energy, commodities, or home prices increase.	N/A	Sentence is not relevant to monetary policy.
RBA Expectations, Actions, Assets	RBA expects subpar inflation, RBA expecting disinflation, narrowing spreads of treasury bonds, decreases in treasury security yields, and reduction of bank reserves.	RBA expects high inflation, widening spreads of treasury bonds, increase in treasury security yields, increase bank reserves.	N/A	Sentence is not relevant to monetary policy.

Category	Dovish	Hawkish	Neutral	Irrelevant
Foreign Nations	When Australia's trade deficit decreases or there are positive developments in trade relations, especially with top trading partners such as China, Japan, US, and the EU.	When Australia's trade deficit increases or trade relations worsen, especially concerning top trading partners such as China, Japan, US, and the EU.	When relating to a foreign nation's policy with no effect.	Sentence is not relevant to monetary policy.
Money Supply	Money supply is low, M2 increases, increased demand for loans.	Money supply is high, increased demand for goods, low demand for loans.	N/A	Sentence is not relevant to monetary policy.
Labor	When productivity increases, when unemployment increases.	When productivity decreases, when unemployment decreases.	N/A	Sentence is not relevant to monetary policy.
Consumer Sentiment	Negative Consumer Sentiment.	Positive Consumer Sentiment.	N/A	Sentence is not relevant to monetary policy.
Wages	Wages decrease.	Wages increase.	N/A	Sentence is not relevant to monetary policy.

2648 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 68 and whether the committee is talking about the sentence with Certainty attest
2649 to it or not in Table 69.
2650

Table 68: Reserve Bank of Australia Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Statements that anticipate or project future economic conditions, expectations, or policy actions. Key words include “outlook,” “expect,” “forecast,” “prospects,” and “likely.”	“The outlook for the labor market remains positive.”
Not Forward Looking	Statements that describe past economic data, trends, or events. Key words include “has been,” “stable,” “last year,” and “past.”	“The economic outlook has been stable for the past year.”

Table 69: Reserve Bank of Australia Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Statements that are definitively stated with clarity and without ambiguity. Key words include “guaranteed,” “remain,” “determined,” “decided,” “mandated,” “confirmed,” and “ensured.”	“The cash rate will remain at 4% until inflation targets are met.”
Uncertain	Statements that express potential risks, conditions, or future possibilities. Key words include “potential,” “may,” “could,” “unclear,” “contingent,” “likely,” and “possibly.”	“The RBA may consider further tightening if inflationary pressures persist.”

2651

Bank of Israel

2652

Region: Israel

2653



2654

2655

Data Collected: 2006-2024

2656

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

2657

2658 Monetary Policy Mandate

2659 The Bank of Israel's main function is to protect the value of the local currency - in other words, to
2660 maintain price stability. Maintaining the value of money is important for economic stability and
2661 strength, and for creating the conditions necessary for continued growth of output and employment.

2662 Mandate Objectives:

- 2663 • **Price Stability:** The Bank of Israel's main function is to protect the value of the local
2664 currency.
- 2665 • **Economic Growth:** Additional functions are to support the attainment of the other goals
2666 of the government's economic policy, especially growth, employment, and the narrowing
2667 of social gaps, as well as the stability and proper activity of the financial system – all these
2668 provided that price stability is not undermined in the long run.

2669 Structure

2670 Composition:

- 2671 • **Officials:** The Governor, who serves as chairperson of the Committee, the Deputy Governor,
2672 and an additional Bank employee who is appointed by the Governor.
- 2673 • **Public Representatives:** The other three members of the Committee are representatives of
2674 the public.

2675 **Meeting Structure:** The Monetary Committee meets eight times a year to discuss the current
2676 economic outlook and makes rate decisions based on the current price stability and economic growth.

2677 Manual Annotation

2678 Annotators

- 2679 • Tyson Brown
- 2680 • Matthew Lu
- 2681 • Gilbert Mao
- 2682 • Daniel Wu

2683 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 2684 • Stance Agreement: 38.7%
- 2685 • Time Agreement: 78.2%
- 2686 • Certainty Agreement: 65.6%

2687 Annotation Guide

2688 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
2689 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
2690 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
2691 other than monetary policy stance.

2692 The annotation guide for monetary policy stance is summarized in Table 70. It is built by dividing
2693 each target sentence into 10 defined categories: Interest Rate, Capital Market Inflation Rate, Total
2694 Assets, Gross Domestic Product (GDP), Unemployment Rate, Job Vacancy Rate, NIS to Euro (or
2695 Dollar), Inflation, Housing Market, War.

2696 Important Factors

- 2697 • *Interest Rate:* The Bank of Israel has adjusted its benchmark interest rate, influencing the
2698 cost of borrowing and impacting economic activity by either stimulating or cooling down
2699 demand.

- 2700 • *Capital Market Inflation Rate*: Inflation in Israel's financial markets has been reflected in
2701 fluctuations in asset prices, with stock and bond markets responding to macroeconomic
2702 conditions and central bank policies.
- 2703 • *Total Assets*: The Bank of Israel's balance sheet has expanded as foreign reserves increased,
2704 injecting liquidity into the financial system and affecting overall economic stability.
- 2705 • *GDP*: Israel's GDP growth has been influenced by the central bank's monetary policies,
2706 balancing inflation control with efforts to sustain economic expansion.
- 2707 • *Unemployment Rate*: The unemployment rate has shifted due to monetary policy measures
2708 aimed at either job creation or inflation containment, affecting labor market dynamics.
- 2709 • *Job Vacancy Rate*: Changes in the job vacancy rate indicate shifts in labor demand, reflecting
2710 the impact of central bank policies on hiring trends and wage pressures.
- 2711 • *NIS to Euro (or Dollar)*: The exchange rate of the Israeli shekel against the Euro and the
2712 U.S. dollar has fluctuated, influencing import costs, export competitiveness, and inflation
2713 levels.
- 2714 • *Inflation*: The Bank of Israel has implemented monetary policies to stabilize inflation,
2715 ensuring that price levels remain within the target range for economic stability.
- 2716 • *Housing Market*: Interest rate adjustments have directly affected mortgage rates, impacting
2717 housing affordability and real estate market activity.
- 2718 • *War*: Ongoing geopolitical conflicts have introduced economic uncertainty, affecting market
2719 confidence, trade flows, and the central bank's policy responses.

2720 **Examples:**

- 2721 • "The Monetary Committee decided to leave the interest rate unchanged, but sees a real
2722 possibility of having to raise the interest rate in future decisions."
2723 **Hawkish**: This sentence implies that despite the current stance, future policy may tighten,
2724 reflecting caution against inflation.
- 2725 • "The Bank of Israel increased bond purchases and foreign currency reserves."
2726 **Hawkish**: This example indicates a proactive measure to curb potential overheating of the
2727 economy.
- 2728 • "The Committee members agreed that if necessary, the Committee will take additional steps
2729 to make monetary policy even more accommodative."
2730 **Dovish**: The sentence suggests a readiness to ease monetary conditions in order to support
2731 economic growth.
- 2732 • "The Committee decided to sell bonds and securities to reduce the money supply."
2733 **Dovish**: By selling bonds, the bank is expanding its balance sheet, a typical sign of an easing
2734 or accommodative policy stance.
- 2735 • "It was decided to keep the interest rate unchanged at 0.1 percent."
2736 **Neutral**: This sentence reflects a steady policy stance with no movement towards tightening
2737 or easing.
- 2738 • "Other participants in the discussion are the directors of the research and market operations
2739 departments, and economists from various departments who prepare and present the material
2740 for discussion."
2741 **Irrelevant**: This sentence discusses the participants in a discussion rather than focusing on
2742 monetary policy or actions that influence monetary policy decisions.

Table 70: Bank of Israel Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Interest Rate	When interest rates are too high, less money to spend, reducing economic growth.	When interest rates are too low, lots of borrowing, reduces value of money.	When interest rates remain the same.	Sentence is not relevant to monetary policy.
Capital Market Inflation Rate	Likely to accept higher inflation for lower rates.	Aim to keep inflation low with tighter policies.	Stable inflation without policy shifts.	Sentence is not relevant to monetary policy.
Total Assets	Expanding balance sheets, such as selling bonds.	Reducing balance sheet to protect against overheating.	Maintaining a stable balance sheet.	Sentence is not relevant to monetary policy.
GDP	Shrinking GDP indicating need for stimulus.	Rising GDP increases inflation/employment.	GDP remains steady.	Sentence is not relevant to monetary policy.
Unemployment Rate	Focus on lowering unemployment.	Willing to tolerate higher unemployment due to inflation.	Balancing unemployment and inflation.	Sentence is not relevant to monetary policy.
Job Vacancy Rate	Stimulating economy to fill vacancies.	High vacancies as wage inflation signal; tighter stance.	Monitored but not policy-shaping.	Sentence is not relevant to monetary policy.
NIS to Euro (or Dollar)	High NIS makes exports expensive, lowering trade.	Higher interest rates to make NIS more attractive.	Inflation at target, balanced economy.	Sentence is not relevant to monetary policy.
Housing Market	Lower interest rates make mortgages more affordable.	Higher rates discourage borrowing, cool prices.	Inflation and rates at target levels.	Sentence is not relevant to monetary policy.
War	Lower risk perception, lower rates encouraged.	Higher risk perception, tighter policy.	War impact unchanged.	Sentence is not relevant to monetary policy.

2743 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 71 and whether the committee is talking about the sentence with Certainty attest
2744 to it or not in Table 72.
2745

Table 71: Bank of Israel Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Discusses expectations, projections, or anticipations about future actions or events. Key words include "future", "expect", "project", "anticipate".	"The Monetary Committee decided to leave the interest rate unchanged, but sees a real possibility of having to raise it in future decisions."
Not Forward Looking	Reflects on recent or past economic data, trends, or events to describe historical or current events. Key words include "past", "historical", "analyzed", "occurred".	"The interest rate can be left at its current level as inflation is expected to converge within the target range beginning in another year."

Table 72: Bank of Israel Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	An expectation, trend, action, or outcome that is clearly stated. Key words include "remain", "increased", "decreased", "will continue".	"As expected, the central bank kept the interest rate unchanged at 0.5 percent this month."
Uncertain	An unclear or unexplained event or action. Key words include "if", "could", "outlook", "uncertainty".	"An open discussion on monetary policy follows, which ends with a vote on the level of the interest rate."

2746

Bank of Canada

2747

Region: Canada

2748



2749

2750

Data Collected: 2005-2024

2751

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

2752

2753 Monetary Policy Mandate

2754 The goal of Canada’s monetary policy is to promote the economic and financial well-being of
2755 Canadians.

2756 Mandate Objectives:

- 2757 • **Controlled Inflation Rate:** Aiming for a low and stable inflation rate at 2% to allow
2758 Canadians to make spending and investment decisions confidently.

2759 Structure

2760 Composition:

2761

- 2762 • **Governing Council:** The Governing Council is the policy-making body of the Bank. It
2763 consists of the Governor, Senior Deputy Governor and the Deputy Governors.
- 2764 • **Monetary Policy Review Committee:** Composed of Governing Council members and
2765 select senior Bank leaders that discuss the challenges and merits of different courses of
2766 action.
- 2767 • **Economics Departments:** Four economics departments, which produce the forecasts and
2768 other analyses that feed into the process

2769 Meeting Structure:

- 2770 • **Frequency:** Eight times per year on fixed dates to decide the policy interest rate.

2771 Manual Annotation

2772 Annotators

- 2773 • Erik Larson
- 2774 • Anish Kanduri
- 2775 • Ryan Sheppard
- 2776 • Ojas Kalra

2777 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 2778 • Stance Agreement: 78%
- 2779 • Time Agreement: 84.8%
- 2780 • Certainty Agreement: 80.4%

2781 **Annotation Guide** Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify
2782 their annotation guide further, keeping in mind the changing global macroeconomic dynamics. Also,
2783 we have two additional categories (Forward Looking or Not Forward Looking and Certain or
2784 Uncertain) of labels other than monetary policy stance.

2785 The annotation guide for monetary policy stance is summarized in Table 73. It is built by dividing
2786 each target sentence into eight defined categories: Inflation, Employment, Economic Growth, Interest
2787 Rates, Monetary Stimulus, Exchange Rates, Commodity Prices, and Fiscal Policy Interaction.

- 2788 • *Inflation:* A sentence pertaining to the evolution of price levels, encompassing trends of
2789 changing inflation rates or pressures on inflation.
- 2790 • *Employment:* A sentence pertaining to labor market conditions, focusing on unemployment
2791 rates, job creation trends, and workforce strength.
- 2792 • *Economic Growth:* A sentence pertaining to the pace and direction of economic expansion
2793 or contraction, including implications for stimulus during slowdowns or overheating during
2794 rapid growth.

- 2795 • *Interest Rates*: A sentence pertaining to interest rate levels, discussing policy decisions
2796 aimed at modifying interest rates to address the needs of the economy.
- 2797 • *Monetary Stimulus*: A sentence pertaining to actions that expand liquidity, such as increased
2798 asset purchases.
- 2799 • *Exchange Rates*: A sentence pertaining to the movement of a nation's currency value or
2800 their economic implications.
- 2801 • *Commodity Prices*: A sentence pertaining to shifts in the prices of Canada's commodities.
- 2802 • *Fiscal Policy Interaction*: A sentence pertaining to the interactions between government
2803 fiscal measures (such as spending and taxation) and monetary policy decisions.

2804 **Examples:**

- 2805 • "Inflation is expected to exceed 3%, requiring tightening."
2806 **Hawkish**: This sentence implies that, due to rising inflation above the target, the Bank of
2807 Canada may adopt tightening policies to control inflationary pressures.
- 2808 • "Unemployment has fallen significantly, prompting concerns about wage inflation."
2809 **Hawkish**: This example indicates that a tight labor market may lead to wage pressures,
2810 suggesting that monetary policy could tighten to prevent an overheating economy.
- 2811 • "The central bank cut rates to support the slowing economy."
2812 **Dovish**: This sentence states that the Bank of Canada took an accommodative stance to
2813 stimulate economic activity during economic slowdowns.
- 2814 • "Weak economic activity signals the need for further monetary stimulus."
2815 **Dovish**: This example reflects the view that subdued growth conditions warrant more easing
2816 to boost economic performance.
- 2817 • "Interest rates remained steady this month."
2818 **Neutral**: This sentence portrays a situation where the policy remains unchanged, reflecting
2819 stability without leaning towards tightening or easing.
- 2820 • "We will take decisions one meeting at a time."
2821 **Irrelevant**: This sentence discusses the way the Bank of Canada plans to make monetary
2822 policy decisions, however, it does not explicitly discuss any monetary policy action.

Table 73: Bank of Canada Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Inflation	Below target or deflation concerns leading to easing.	Above target or rising inflation, prompting tightening.	Descriptive with no policy implication.	Sentence is not relevant to monetary policy.
Employment	High unemployment, weak labor market prompting easing.	Low unemployment, risk of overheating, tightening bias.	Labor description without policy implication.	Sentence is not relevant to monetary policy.
Economic Growth	Slow growth or recession risk, stimulus needed.	Strong growth, risk of overheating.	Growth data with no implied action.	Sentence is not relevant to monetary policy.
Interest Rates	Lowering or keeping rates low to support economy.	Raising rates to counter inflation.	Rate level mentioned without bias.	Sentence is not relevant to monetary policy.
Monetary Stimulus	Increasing liquidity or asset purchases.	Reducing purchases or liquidity, signaling tightening.	Describes programs neutrally.	Sentence is not relevant to monetary policy.
Exchange Rates	Depreciation prompts easing to support growth.	Appreciation prompts concern about overheating.	Neutral exchange rate statement.	Sentence is not relevant to monetary policy.
Commodity Prices	Falling prices suggest weak demand, needing easing.	Rising prices suggest inflation, tightening needed.	Price movement with no policy stance.	Sentence is not relevant to monetary policy.
Fiscal Policy Interaction	Monetary policy needed to support lagging fiscal stimulus.	Strong fiscal stimulus might require tightening.	Neutral fiscal commentary.	Sentence is not relevant to monetary policy.

2823 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 74 and whether the committee is talking about the sentence with Certainty attest
2824 to it or not in Table 75.
2825

Table 74: Bank of Canada Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Sentences that predict future economic conditions, policy actions, or outcomes. Key phrases include “expected to,” “forecast,” “anticipated to occur.”	“Government spending is not projected to contribute to growth in 2012.”
Not Forward Looking	Sentences referencing past or present economic data, actions, or outcomes. These include descriptions of historical economic data, past policy actions, or statement discussing recent trends.	“Recent data show the Canadian economy is returning to potential growth.”

Table 75: Bank of Canada Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Sentences presenting definitive statements or outcomes with high confidence. Key phrases include “is,” “has been,” “will be.”	“Inflation expectations remain well-anchored.”
Uncertain	Sentences suggesting variability, unpredictability, or tentative outcomes. Key phrases include “could,” “might,” “expected,” “may.”	“Economic growth might slow next quarter depending on global market conditions.”

2826

Bank of Mexico

2827

Region: Mexico

2828



2829

2830

Data Collected: 2018-2024

2831

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

2832

2833 Monetary Policy Mandate

2834 The Banco de Mexico aims to control the value of Mexico's currency and improve the wellbeing of
2835 its citizens.

2836 Mandate Objectives:

- 2837 • **Stable currency rate:** The main mandate of the Banco de Mexico is to maintain a stable
2838 value for the value of Mexico's currency in the long term.

2839 Structure

2840 Composition:

- 2841 • **Governor:** There is one governor on the Board, appointed by the President and confirmed
2842 by the Senate. The Governor has a six-year term, starting on the first of January of the fourth
2843 year of each presidential term.
- 2844 • **Deputy Governors:** There are four deputy governors in the board, appointed by the
2845 President and confirmed by the Senate. The Deputy Governors have an eight-year period.
2846 To be replaced or reappointed every two years, on the first, third, and fifth years of each
2847 presidential term.

2848 Meeting Structure:

- 2849 • **Frequency:** The Board of Governors meets eight times annually on prefixed dates.

2850 Manual Annotation

2851 Annotators

- 2852 • Arnav Tripathi
- 2853 • Chittebbayi Penugonda
- 2854 • Christian Lindler
- 2855 • Surya Subramanian

2856 Annotation Agreement

2857 The agreement percentage among the pairs of annotators for different labels.

- 2858 • Stance Agreement: 57.5%
- 2859 • Time Agreement: 77.8%
- 2860 • Certainty Agreement: 63.3%

2861 Annotation Guide

2862 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
2863 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
2864 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
2865 other than monetary policy stance.

2866 The annotation guide for monetary policy stance is summarized in Table 76. It is built by dividing
2867 each target sentence into 7 defined categories: Peso Value Change, Trade Activity, Labor, Key Words,
2868 Economic Growth, Market Sentiment, Energy/House Prices.

- 2869 • *Peso Value Change:* Sentences corresponding to peso value changes such as inflation and/or
2870 foreign market appreciation/depreciation. A weaker peso may be seen as dovish, especially
2871 when Banco de México is attempting to support exports by making Mexican goods cheaper
2872 in foreign markets. A stronger peso helps control inflation by making imports cheaper, but it
2873 can hurt export competitiveness.

- 2874 • *Trade Activity:* Sentences pertaining to the import/export ratio—an increase in one or
2875 decrease of another. Banco de México favors policies that stimulate exports, especially
2876 given Mexico’s strong manufacturing sector and trade agreements (such as USMCA). A
2877 focus on export growth could support job creation and help maintain balance in current
2878 account deficits.
- 2879 • *Labor:* Sentences mentioning statistics about the increase or decrease in productivity. If
2880 there are concerns about declining productivity—especially in key industries in Mexico like
2881 oil, manufacturing, or agriculture—Banco de México might lean hawkish, signaling the
2882 need to cool off the economy to address inefficiencies.
- 2883 • *Key Words:* Sentences that have keywords that make a clear, strong indication towards
2884 classifications.
- 2885 • *Economic Growth:* Sentences conveying the performance in economic growth the bank
2886 wishes to see in the near future. Given Mexico’s economic structure, fostering growth in
2887 sectors like manufacturing, tourism, and remittances can be seen as a dovish move aimed at
2888 reducing poverty and unemployment.
- 2889 • *Market Sentiment:* Sentences mentioning what the outlook is for the market and how
2890 investors are acting.
- 2891 • *Energy/House Prices:* Sentences that mention anything about housing price levels and
2892 energy price levels. Given Mexico’s reliance on oil and gas revenues, a dovish outlook
2893 might stem from efforts to lower energy prices domestically, encouraging consumption and
2894 reducing inflationary pressure on households.

2895 **Examples:**

- 2896 • “Another member noted that non-core inflation has been below forecasts.”
2897 **Dovish:** This sentence states that non-core inflation is lower than expected, possibly opening
2898 up opportunities for expansionary monetary policy.
- 2899 • “However, another member specified that breakeven inflation still remains at levels above
2900 4% for all terms.”
2901 **Hawkish:** This sentence highlights that breakeven inflation exceeds 4% across all terms,
2902 hinting at potential future tightening to counter inflation concerns.
- 2903 • “In this context, some members noted that the confidence of economic agents decreased.”
2904 **Dovish:** This sentence suggests that a decline in economic confidence might encourage a
2905 more supportive monetary policy approach to improve confidence.
- 2906 • “Another member stated that the dynamics within the components of inflation has led to an
2907 upward adjustment of forecasts.”
2908 **Hawkish:** This sentence asserts that specific dynamics have already resulted in a revision
2909 upward of forecasts, requiring contractionary monetary policy to address inflation concerns.
- 2910 • “Some members considered that the economy’s cyclical position have not changed signifi-
2911 cantly since the previous monetary policy decision.”
2912 **Neutral:** This sentence indicates that the economy’s cyclical stance remains stable, implying
2913 no immediate change in policy direction.
- 2914 • “One member added that the bank of Canada is also preparing to do so.”
2915 **Irrelevant:** This sentence refers to potential preparatory actions by the Bank of Canada
2916 without directly addressing a monetary policy decision or stance.

Table 76: Bank of Mexico Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Peso Value Change	Peso appreciates, leading to lower inflation.	Peso depreciates, leading to higher inflation.	N/A	Sentence is not relevant to monetary policy.
Trade Activity	Favors decreased imports and increased exports, supporting domestic growth.	Favors increased imports and decreased exports, raising dependency on foreign markets.	N/A	Sentence is not relevant to monetary policy.
Labor	Favors increased labor productivity, improving economic output.	Favors lower productivity, signaling economic slowdown.	Labor market remains stable, no notable changes.	Sentence is not relevant to monetary policy.
Key Words	Words like “Accommodative”, “Expansionary”, “Easing”, indicating support for stimulus.	Words like “Tighter”, “Restrictive”, “Balance”, indicating a shift toward tightening.	Words like “Mixed”, “Moderate”, signaling a neutral stance.	Sentence is not relevant to monetary policy.
Economic Growth	Supports expansion, higher consumer spending, and investment.	Advocates policies leading to economic slowdown or controlled growth.	Economy remains balanced, no significant shifts in output.	Sentence is not relevant to monetary policy.
Market Sentiment	Bullish outlook, rising investor confidence in the economy.	Bearish outlook, increased caution among investors.	Neutral market stance, adopting a wait-and-see approach.	Sentence is not relevant to monetary policy.
Energy/House Prices	Rising property values and lower energy costs, benefiting consumers.	Housing market slowdown, rising energy costs, increasing financial burden.	N/A	Sentence is not relevant to monetary policy.

2917 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 77 and whether the committee is talking about the sentence with Certainty attest
2918 to it or not in Table 78.
2919

Table 77: Bank of Mexico Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Focusing on future. Key words include “expect,” “anticipate,” and “project.”	“We expect inflation to decline over the next quarter.”
Not Forward Looking	Focusing on past. Key words include “remained” and “over past X years/months.”	“Inflation remained stable over the past year.”

Table 78: Bank of Mexico Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Sentiment is sure. Key words include “reported,” “show,” and “according to.”	“Reports confirm that economic growth accelerated this quarter.”
Uncertain	Sentiment is unsure. Key words include “likely,” “estimate,” and “projected.”	“It is estimated that inflation may rise in the next period.”

2920
2921
2922

2923
2924

Narodowy Bank Polski

Region: Poland



Data Collected: 2007-2024

2925
2926

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

2927 Monetary Policy Mandate

2928 The MPC is responsible for formulating Poland's monetary policy to achieve the primary object of
2929 maintaining price stability.

2930 Mandate Objectives:

- 2931 • **Price Stability:** Aiming for a 2.5% annual inflation rate, as measured by annual change in
2932 the consumer price index, with a symmetric band for deviations of $\pm 1\%$ for the medium
2933 term.

2934 Structure

2935 The MPC is integral to the National Bank of Poland as they set the economy's monetary policy by
2936 influencing interest rates and the money supply.

2937 Composition:

- 2938 • **Chairperson:** The Governor of the NBP. They are appointed by the Sejm (the lower house
2939 of Poland's parliament) and serve a 6-year term.
- 2940 • **Committee Members:** Nine members appointed in equal proportions by the President,
2941 Senate, and the Sejm. They serve a 6-year term.

2942 Meeting Structure:

- 2943 • **Frequency:** The MPC typically meets once a month, with each meeting lasting two days.
- 2944 • **Non-decision-making Meetings:** The MPC also holds non-decision-making meetings to
2945 analyze the economy and explore policy options.
- 2946 • **Additional Meetings:** Held as needed to address urgent economic developments.

2947 Manual Annotation

2948 Annotators

- 2949 • Saketh Budideti
- 2950 • Huzaifa Pardawala
- 2951 • Harsit Mittal
- 2952 • Rachel Yuh

2953 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 2954 • Stance Agreement: 75.8%
- 2955 • Time Agreement: 86%
- 2956 • Certainty Agreement: 81.3%

2957 **Annotation Guide** Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify
2958 their annotation guide further, keeping in mind the changing global macroeconomic dynamics. Also,
2959 we have two additional categories (Forward Looking or Not Forward Looking and Certain or
2960 Uncertain) of labels other than monetary policy stance.

2961 The annotation guide for monetary policy stance is summarized in Table 79. It is built by dividing each
2962 target sentence into 13 defined categories: Interest Rates, Inflation Rate, GDP Growth, Unemployment
2963 Rate, Required Reserve Ratio, Money Supply, Foreign Exchange Reserves, Private Consumption,
2964 Prices, Lending, Domestic Manufacturing, Gold Holdings, Net Exports.

- 2965 • *Interest Rates:* A sentence pertaining to various interest rates used by the NBP. This includes,
2966 but is not limited to, the reference rate, Lombard rate, deposit rate, discount rate, and
2967 inflation rate.

- 2968 • *Inflation Rate*: A sentence pertaining to the inflation rate (the change in the consumer price
2969 index).
- 2970 • *GDP Growth*: A sentence pertaining to the GDP growth of Poland.
- 2971 • *Unemployment Rate*: A sentence pertaining to the unemployment rate of Poland.
- 2972 • *Required Reserve Ratio*: A sentence pertaining to the required reserve ratio, that is, the
2973 minimum percentage of reservable liabilities that depository institutions (e.g. banks) must
2974 hold at any given time.
- 2975 • *Money Supply*: A sentence that discusses the impact to money supply or changes in demand.
- 2976 • *Foreign Exchange Reserves*: A sentence pertaining to foreign assets controlled by the NBP.
- 2977 • *Private Consumption*: A sentence pertaining to private consumption in Poland.
- 2978 • *Prices*: A sentence pertaining to a change in prices of goods and services in Poland.
- 2979 • *Lending*: A sentence that discusses changes in lending behaviors of actors such as corpora-
2980 tions.
- 2981 • *Domestic Manufacturing*: A sentence discusses the growth or output of the manufacturing
2982 industry in Poland.
- 2983 • *Gold Holdings*: A sentence pertaining to the gold holdings of the NBP.
- 2984 • *Net Exports*: A sentence discussion the net exports of Poland.

2985 **Examples:**

- 2986 • “The Council judged that the current level of the NBP interest rates was conducive to meeting
2987 the NBP inflation target in the medium term”
2988 **Neutral**: Interested rates are expected to meeting the inflation target, signaling that no
2989 change should be made.
- 2990 • “Some Council members pointed to a continually high rate of wage growth and to a very
2991 fast drop in unemployment, facilitating a further strong rise in wages.”
2992 **Hawkish**: High wage growth and low unemployment indicate wage-driven inflation, requir-
2993 ing tighter monetary policy to combat.
- 2994 • “Some Council members stressed that in the event of significant price growth that would
2995 jeopardize meeting the inflation target in the medium term, it might be justified to consider
2996 an increase in the NBP interest rates in the coming quarters.”
2997 **Hawkish**: The emphasis on potential interest rates increases to address price growth signals
2998 commitment to tighten monetary policy amid inflation risks.
- 2999 • “These meeting participants also assessed that the acceleration of interest rate increases
3000 might not reduce inflation expectations of households, which are highly adaptive in nature.”
3001 **Dovish**: Participants questions the effectiveness of rate hikes in lowering inflation expecta-
3002 tions, suggesting a looser monetary policy stance.
- 3003 • “In effect, it was judged that the annual GDP growth in 2022 Q4 slowed down markedly for
3004 another consecutive quarter.”
3005 **Dovish**: Slowing GDP growth indicates a need for easing to stimulate economic momentum.
- 3006 • “The discussion at the meeting focused on the outlook for economic growth abroad and in
3007 Poland, fiscal policy, zloty exchange rate developments and the situation in the credit market
3008 and the banking sector.”
3009 **Irrelevant**: The sentence discusses the topics of the meeting, indicating no monetary policy
3010 stance.

Table 79: Narodowy Bank Polski Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Interest Rates	When interest rates decrease.	When interest rates increase.	When interest rates are unchanged, maintained, or sustained.	Sentence is not relevant to monetary policy.
Inflation Rate	When inflation is below target.	When inflation is above target.	When inflation is at target.	Sentence is not relevant to monetary policy.
GDP Growth	Slowing GDP growth.	When GDP growth is at a sustained increase above potential.	Stable GDP growth.	Sentence is not relevant to monetary policy.
Unemployment Rate	Increase in unemployment rates.	Decrease in unemployment rates.	Stable unemployment rates.	Sentence is not relevant to monetary policy.
Required Reserve Ratio	Decrease in required reserve ratio.	Increase in required reserve ratio.	Maintaining the required reserve ratio.	Sentence is not relevant to monetary policy.
Money Supply	Low money supply or increased demand for loans.	Money supply is high, increased demand for goods, low demand for loans.	Money supply is stable.	Sentence is not relevant to monetary policy.
Foreign Exchange Reserves	Increasing reserves, indicating boosting market confidence.	Decreasing reserves.	Maintained reserves.	Sentence is not relevant to monetary policy.
Private Consumption	Decrease in consumption due to weak consumer sentiment.	Increase in consumption due to rising wages or government spending.	Sustained private consumption.	Sentence is not relevant to monetary policy.
Prices	Falling prices indicating reduced inflation.	Rising prices indicating increased inflation.	Maintaining existing price levels.	Sentence is not relevant to monetary policy.
Corporate Lending	Decline in lending, low investment.	Rapid increase in corporate borrowing.	Maintaining existing lending levels.	Sentence is not relevant to monetary policy.
Domestic Manufacturing	Weak domestic manufacturing output or growth.	Significant domestic manufacturing output or growth.	Stable domestic manufacturing output and growth.	Sentence is not relevant to monetary policy.
Gold Holdings	Falling gold holdings indicating flexibility in monetary policy.	Rising gold holdings indicating tighter monetary policy.	Maintained gold holdings.	Sentence is not relevant to monetary policy.

Category		Hawkish	Dovish	Neutral	Irrelevant
Net	Ex-	Low net exports indicative of low demand or uncompetitive industries.	High net exports, indicating higher risk of inflation.	Maintained net export levels.	Sentence is not relevant to monetary policy.

3011 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 80 and whether the committee is talking about the sentence with Certainty attest
3012 to it or not in Table 81.
3013

Table 80: Narodowy Bank Polski Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	text describing events anticipated to occur in the future. Sentences often utilize language such as “expected” or “projected.”	“The Council underscored that data incoming in 2023 Q3 signalled a further deterioration in economic conditions in the euro area, including in Germany.”
Not Forward Looking	text referring to past or recent economic policy, data, trends, or events.	“In this context, it was noted that corporate lending, including for investments, had rebounded.”

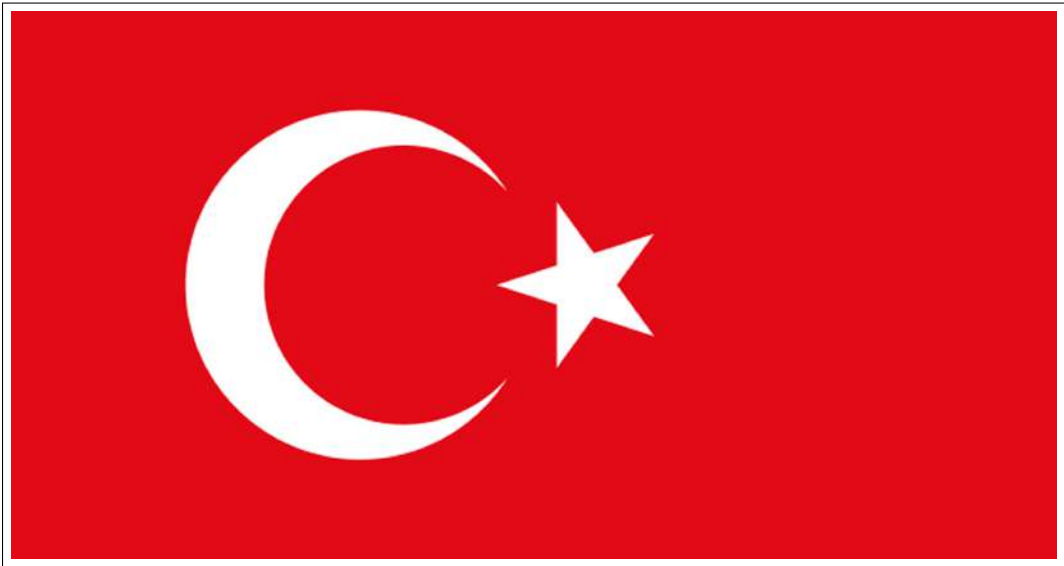
Table 81: Narodowy Bank Polski Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	The sentence reflects a clear, firm, or definitive decision or expectation about actions or outcomes without ambiguity.	“The Council decided to increase the NBP interest rates to the following levels reference rate to 4.75, lombard rate to 6.25, deposit rate to 3.25, rediscount rate to 5.00.”
Uncertain	The sentence expresses potential risks, varying opinions, or uncertainties about events, outcomes, or actions, often using words like “could,” “might,” or “may.”	“These meeting participants also assessed that the acceleration of interest rate increases might not reduce inflation expectations of households, which are highly adaptive in nature.”

3014
3015
3016
3017
3018

Central Bank of Turkey

Region: Turkey



Data Collected: 2015-2024

3019
3020

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

3021 Monetary Policy Mandate

3022 The Central Republic Bank of Turkey (CBRT) is responsible for formulating Türkiye’s monetary
3023 policy to achieve the primary objective of maintaining price stability.

3024 Mandate Objectives:

- 3025 • **Price Stability:** The primary goal of the CBRT is to achieve and maintain price stability,
3026 using an inflation targeting regime to guide monetary policy.
- 3027 • **Inflation Targeting:** Since 2006, the CBRT has implemented an explicit inflation targeting
3028 framework, setting annual inflation targets in coordination with the government, with an
3029 uncertainty band to account for fluctuations.
- 3030 • **Financial Stability:** Post-2010, the CBRT enhanced its monetary policy to also safeguard fi-
3031 nancial stability, introducing tools like required reserves and the Reserve Options Mechanism
3032 to mitigate external shocks.

3033 Structure

3034 The Monetary Policy Committee (MPC) is responsible for setting the economy’s monetary policy by
3035 determining interest rates and guiding liquidity management, making it key to the CRBT.

3036 Composition:

- 3037 • **Governor:** The Governor of the CBRT, who also serves as the chair of the MPC. The
3038 Governor is appointed by the President of Türkiye.
- 3039 • **Committee Members:** The MPC consists of the Governor, the Deputy Governors of the
3040 CBRT, and additional members appointed by the President.

3041 Meeting Structure:

- 3042 • **Frequency:** At least eight times a year according to a pre-announced schedule.
- 3043 • **Regular Reports:** The CBRT publishes a Monthly Price Developments Report following
3044 inflation data releases and a Financial Stability Report analyzing the financial system.
- 3045 • **Non decision-making meetings:** The committee holds discussions on economic develop-
3046 ments and policy options outside of formal decision-making sessions.

3047 Manual Annotation

3048 Annotators

- 3049 • Veer Guda
- 3050 • Tarun Mandapati
- 3051 • Tamas Mester
- 3052 • Alexander Wang

3053 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 3054 • Stance Agreement: 54.2%
- 3055 • Time Agreement: 88.2%
- 3056 • Certainty Agreement: 89.5%

3057 Annotation Guide

3058 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
3059 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
3060 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
3061 other than monetary policy stance.

3062 The annotation guide for monetary policy stance is summarized in Table 82. It is built by dividing
 3063 each target sentence into 6 defined categories: Economic Status, Currency Exchange Rate (Turkish
 3064 Lira/TRY), Inflation Rates, Interest Rates, Government Debt & Fiscal Policy, and Foreign Trade
 3065 Balance.

- 3066 • *Economic Status*: Represents the overall condition of the Turkish economy, including
 3067 inflation, unemployment, and growth trends.
- 3068 • *Currency Exchange Rate (Turkish Lira - TRY)*: Fluctuations in the value of the Turkish lira
 3069 relative to other currencies, influencing trade and inflation.
- 3070 • *Inflation Rates*: Measures the rate of price increases for goods and services in Turkey,
 3071 impacting purchasing power and policy decisions.
- 3072 • *Interest Rates*: The CBRT's benchmark policy rate, affecting borrowing costs, investment,
 3073 and overall economic activity.
- 3074 • *Government Debt & Fiscal Policy*: Represents Turkey's government borrowing levels and
 3075 fiscal management, influencing monetary policy choices.
- 3076 • *Foreign Trade Balance*: Measures Turkey's exports and imports, with trade surpluses or
 3077 deficits impacting currency strength and policy decisions.

3078 Examples:

- 3079 • "Thus, the unemployment rate declined by 0.4 points compared to previous month and
 3080 realized as 9.6%."
 3081 **Hawkish**: The sentence reports a decline in unemployment, potentially justifying a more
 3082 hawkish approach to curb inflation.
- 3083 • "Thus, the rise in automobile prices amounted to 34.01% in the first six months of 2023."
 3084 **Hawkish**: This sentence highlights a significant increase in automobile prices, suggesting
 3085 potential inflationary pressures that may prompt a tighter monetary policy stance.
- 3086 • "The committee has therefore decided to keep the interest rates at current levels."
 3087 **Neutral**: The sentence simply states that interest rates will remain unchanged, reflecting no
 3088 immediate bias toward tightening or easing policy.
- 3089 • "Annual inflation remained flat in rents, but decreased in other subgroups."
 3090 **Dovish**: With inflation showing stability in rents and declines elsewhere, the overall price
 3091 pressure appears subdued, which may support expansionary monetary policy.
- 3092 • "In sum, economic activity is expected to recede in the third quarter."
 3093 **Dovish**: This forward Looking forecast of a slowdown in economic activity points to the
 3094 possibility of an easing monetary stance to promote economic growth.
- 3095 • "Impact analyses of the CBRT's regulations are conducted for all components of the frame-
 3096 work by assessing their impact on inflation, interest rates, exchange rates, reserves, expecta-
 3097 tions, securities, and financial stability with a holistic approach."
 3098 **Irrelevant**: The sentence focuses on a comprehensive evaluation process of regulatory
 3099 impacts without providing any specific insight into monetary policy stance.

Table 82: Central Bank of Turkey Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Economic Status	When inflation is decreasing, unemployment is rising, or economic growth is weak.	When inflation is increasing, unemployment is low, and economic growth is strong.	When economic indicators remain stable.	Sentence is not relevant to monetary policy.
Currency Exchange Rate (Turkish Lira/TRY)	When the lira depreciates significantly, increasing import costs and inflationary pressure.	When the lira appreciates, reducing import costs and signaling potential monetary tightening.	When the exchange rate remains stable.	Sentence is not relevant to monetary policy.
Inflation Rate	When inflation is slowing down, suggesting room for rate cuts.	When inflation is rising, requiring tighter monetary policy.	When inflation remains within the target range.	Sentence is not relevant to monetary policy.
Money Supply	Low money supply, any increase in M2, demand for loans is higher.	High money supply, demand for goods increase, demand for loans decrease.	N/A.	Sentence is not relevant to monetary policy.
Government Debt	When government debt increases significantly, prompting expectations of monetary stimulus to manage borrowing costs.	When debt decreases, indicating fiscal discipline and allowing tighter monetary policy.	When debt levels remain stable.	Sentence is not relevant to monetary policy.
Foreign Trade Balance	When the trade deficit widens, indicating higher imports than exports, potentially justifying a weaker lira and looser monetary policy.	When the trade deficit narrows or a surplus is recorded, strengthening the lira and allowing for tighter policy.	When the trade balance remains relatively unchanged.	Sentence is not relevant to monetary policy.
Key Words and Phrases	When the sentiment is growth oriented, e.g., "accommodative," "maximum employment," "price stability," "heightened consumption," "low inflation."	When sentiment is restrictive, indicating a focus on "price stability" and "sustained growth."	Use of neutral terms, e.g., "mixed," "moderate," "reaffirmed."	Sentence is not relevant to monetary policy.

3100 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 83 and whether the committee is talking about the sentence with Certainty attest
3101 to it or not in Table 84.
3102

Table 83: Central Bank of Turkey Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Statements describing anticipated economic developments, monetary policy actions, or financial trends. Often includes phrases like “expected,” “projected,” or “forecasted.”	“The CBRT expects inflation to moderate in the second half of the year as the impact of recent monetary tightening measures takes hold.”
Not Forward Looking	Statements referring to past or current economic trends, monetary decisions, or historical data points.	“Inflation in the previous quarter rose to 70.3%, largely driven by food and energy price hikes.”

Table 84: Central Bank of Turkey Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	The sentence expresses a firm, definitive decision or expectation about monetary policy, economic trends, or financial markets. Often avoids speculative language.	“The CBRT has decided to raise the policy rate from 42.5% to 45% to counter inflationary pressures.”
Uncertain	The sentence reflects potential risks, varying opinions, or uncertainties about future monetary policy or economic conditions, using words such as “may,” “could,” or “might.”	“The Monetary Policy Committee noted that further rate hikes might be necessary if inflation does not show signs of stabilizing.”

3103
3104
3105

3106
3107

Bank of Thailand

Region: Thailand



Data Collected: 2011-2024

3108
3109

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

3110 **Monetary Policy Mandate**

3111 The Bank of Thailand (BoT) aims to foster a stable, sustainable and inclusive macroeconomic and
3112 financial environment in order to allow Thai citizens to sustainably achieve higher quality of life over
3113 the long term.

3114 **Mandate Objectives:**

- 3115 • **Quality of Life:** BoT focuses heavily on the quality of life of the Thai citizens.
- 3116 • **Stable financial environment:** The bank aims to have a sustainable, stable and inclusive
- 3117 macroeconomic and financial environment.

3118 **Structure**

3119 The Bank of Thailand has the following meeting composition and structure:

3120 **Composition:** The MPC will continue to comprise of 7 members, as in the present arrangement.
3121 The BOT Governor will act as Chairman, given the Governor's responsibility in implementing the
3122 MPC's decisions. The Governor also acts as a linkage with the other policy committees of the BOT.
3123 Two deputy Governors of the BOT and four external members form the remaining membership of the
3124 MPC, who aim to provide a balance of views, as intended in the new BOT Act. External members
3125 have a fixed term of 3 years, and cannot serve more than two consecutive terms.

- 3126 • **Internal Members:** There will be 3 internal members. The BOT Governor will act as
- 3127 Chairman, given the Governor's responsibility in implementing the MPC's decisions. The
- 3128 Governor also acts as a linkage with the other policy committees of the BOT. Two deputy
- 3129 Governors of the BOT are the other internal members.

- 3130 • **External Members** Four external members form the remaining membership of the MPC,
- 3131 who aim to provide a balance of views, as intended in the new BOT Act. External members
- 3132 have a fixed term of 3 years, and cannot serve more than two consecutive terms.
- 3133

3134 **Meeting Structure:**

- 3135 • **Frequency:** Six scheduled meetings annually.
- 3136 • **Additional Meetings:** Held as needed.

3137 **Manual Annotation**

3138 **Annotators**

- 3139 • Eric Liao
- 3140 • Tyler Huang
- 3141 • Kenny Cao
- 3142 • Christopher Traylor

3143 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 3144 • Stance Agreement: 35.5%
- 3145 • Time Agreement: 59.1%
- 3146 • Certainty Agreement: 60.7%

3147 **Annotation Guide**

3148 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
3149 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
3150 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
3151 other than monetary policy stance.

3152 The annotation guide for monetary policy stance is summarized in Table 85. It is built by dividing
 3153 each target sentence into 9 defined categories: Inflation, Economic Growth, Copper Prices, Peso
 3154 Exchange Rate, Foreign Reserves, External Demand (Trade Partners), Key Words/Phrases, Labor
 3155 Market, and Energy Prices.

- 3156 • *Economic Status*: A sentence pertaining to the state of the economy, relating to factors such
 3157 as unemployment and inflation.
- 3158 • *Thai Baht Value Change*: A sentence pertaining to the appreciation or depreciation of the
 3159 Thai Baht in the foreign exchange market.
- 3160 • *Energy/House Prices*: A sentence pertaining to fluctuations in energy sectors or housing
 3161 prices, indicative of broader economic conditions.
- 3162 • *Foreign Policy*: A sentence pertaining to trade relations and economic interactions between
 3163 Thailand and its international partners. If not directly related, the sentence is labeled neutral.
- 3164 • *MPC Expectations/Actions/ Assets*: A sentence that discusses the MPC's projections, actions,
 3165 or asset valuations, influencing monetary policy decisions.
- 3166 • *Money Supply*: A sentence that overtly discusses changes in the money supply or shifts in
 3167 the demand for loans.
- 3168 • *Key Words/Phrases*: A sentence that encapsulates specific key words or phrases which
 3169 classify monetary policy stances based on their conventional usage.
- 3170 • *Labor Market*: A sentence pertaining to changes in labor market performance and productiv-
 3171 ity that may affect monetary policy.

3172 Examples:

- 3173 • "The latest risk-adjusted inflation forecast for 2024 eased to 3.9 percent from 4.2 percent in
 3174 the previous meeting in December."
 3175 **Dovish**: The sentence reports a decline in inflation forecast, which reduces the urgency for
 3176 tightening of monetary policy and suggest less inflationary pressure.
- 3177 • "Headline inflation increased in June and July driven mainly by rising transport costs such
 3178 as higher domestic petroleum prices and tricycle fare hikes."
 3179 **Hawkish**: This sentence highlights an increase in headline inflation, suggesting potential
 3180 inflationary pressures that may prompt a tighter monetary policy stance.
- 3181 • "The highlights of the discussions on the 1st December 2011 meeting were approved by the
 3182 monetary board during its regular meeting held on 15th Decemeber 2011."
 3183 **Irrelevant**: The sentence is simply procedural and does not provide any information about
 3184 the monetary policy stance.
- 3185 • "Nonetheless, inflation forecasts for 2016-2017 stayed close to the midpoint of the target
 3186 range."
 3187 **Neutral**: The sentence indicates that the inflation forecasts remained near the midpoint of
 3188 the target range, it implies that inflation is stable but does not convey any information on
 3189 need for monetary tightening or easing.
- 3190 • "However, the moderation in growth in Q1 2014, particularly in the US, as well as tighter
 3191 financial conditions and geopolitical tensions in the Middle East and Russia, could dampen
 3192 the overall momentum for the rest of 2014."
 3193 **Dovish**: This sentence indicates some downside risks to economic growth, which hints at
 3194 potential easing measures.

Table 85: Bank of Thailand Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Economic Status	When inflation decreases, When unemployment increases, when economic growth is projected as low.	When inflation increases, when unemployment decreases when economic growth is projected high when economic output is higher than potential supply/actual output when economic slack falls.	When unemployment rate or growth is unchanged, maintained, or sustained.	Sentence is not relevant to monetary policy.
Thai Baht Value Change	When the Thai Baht appreciates.	When the Thai Baht depreciates.	When there is no real change in Thai Baht worth.	Sentence is not relevant to monetary policy.
Energy/House Prices	When oil/energy sectors dwindle and prices decrease, when housing prices decrease because this can indicate a smaller consumer output, and with a dovish economy there will be a larger consumer output.	When oil/energy sectors are augmented/prices increase, when housing prices increase because this can indicate a large consumer output, and the economy could benefit from a hawkish policy as it will help stabilize prices.	When there is market uncertainty in the oil/energy sectors, when housing prices have an uncertain outlook as this means there is not clear indicators on how consumer spending is.	Sentence is not relevant to monetary policy.
Foreign Policy	When partner economics are starting to expand, Thailand's monetary policy will be more free as Thailand's economic output throughput is directly proportional to its trading partners.	When partner economics are starting to fall or become highly volatile, Thailand should be more hawkish as the partner trading profitability is uncertain.	When describing specifics on an unrelated foreign nation's economic or trade policy as this is likely to have little effect on the monetary policy control for Thailand or uncertainty in partners economies.	Sentence is not relevant to monetary policy.

Category	Dovish	Hawkish	Neutral	Irrelevant
MPC Expectations/ Actions/ Assets	MPC expects subpar inflation, disinflation, narrowing of spread in Thai government bonds/lower yields, low debt serviceability because in these cases, it is in best interest to allow a freer monetary policy to boost spending/market growth.	MPC expects high inflation, widening of spread in Thai government bonds/higher yields, high debt serviceability because in these cases, the market has some room in which the monetary policy may be more strict without a huge negative effect.	MPC's expectations of future economic status remainings uncertain because there is not much indicator for which way the economic status should sway.	Sentence is not relevant to monetary policy.
Money Supply	Money supply is low, M2 measure increases, increased demand for loans.	Money supply is high, increased demand for goods, low demand for loans.	N/A.	Sentence is not relevant to monetary policy.
Key Words and Phrases	When the stance is "accommodative", indicating a focus on "maximum employment" and "price stability".	Indicating a focus on "price stability" and "sustained growth".	Use of phrases "mixed", "moderate", "reaffirmed.	Sentence is not relevant to monetary policy.
Labor Market	When labor markets see improvements, Thailand's policy is more dovish because it leads to lower inflationary pressure to support economic growth and employment.	When labor markets see a decline, Thailand's policy can be more hawkish because it can increase inflationary pressures due to more limited supplies.	When labor markets don't see a change in the productivity of labor markets, there isn't a change in the monetary policy.	Sentence is not relevant to monetary policy.

3195 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 86 and whether the committee is talking about the sentence with Certainty attest
3196 to it or not in Table 87.
3197

Table 86: Bank of Thailand Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Sentences that reflect projections or expectations about future economic conditions. They highlight anticipated trends such as inflation moderation or predicted changes in unemployment.	“Inflation is anticipated to moderate in the coming year.”
Not Forward Looking	Sentences that refer to past economic performance or recent data points, such as prior inflation rates or past GDP growth.	“Last quarter’s GDP growth was lower than expected.”

Table 87: Bank of Thailand Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Sentences presenting definitive statements with high confidence. These include clear projections on economic growth or unambiguous stances on monetary policy.	“The central bank will raise interest rates in the next meeting.”
Uncertain	Sentences that incorporate words or phrases implying doubt or unpredictability, such as “may,” “could,” or “is expected to.”	“Economic growth might slow amid global economic uncertainties.”

3198

Central Bank of Egypt

3199

Region: Egypt

3200



3201

3202

Data Collected: 2012-2024

3203

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

3204

3205 Monetary Policy Mandate

3206 The Monetary Policy Committee (MPC) oversees the Central Bank of Egypt and is responsible for
3207 setting price targets, overnight interbank rate, and other policy rates.
3208

3209 Mandate Objectives:

- 3210 • **Minimize the Inflation Gap:** Transitioning the economy to a “flexible inflation-targeting
3211 framework” by setting several inflation targets to transition from the current inflation level
3212 to medium-term level inflation. Goal was to achieve 7% ($\pm 2\%$) by Q4 2024 5% ($\pm 2\%$) by
3213 Q4 2026. Currently core inflation is 23.220% y/y.
- 3214 • **Minimize the Output Gap:** Striving the maximize output with respect of full capacity
3215 maximization while minimizing macroeconomic volatility and maintaining price stability

3216 Structure

3217 **Composition:** The MPC consists of a Governor, two deputy Governors, six members, and three
3218 non-executive Board Members, selected by the Board of Directors.

3219 Meeting Structure:

- 3220 • **Frequency:** Eight times a year, the schedule of which is announced at the beginning of the
3221 year
- 3222 • **Additional Meetings:** Held as necessary in extraordinary circumstances

3223 Manual Annotation

3224 Annotators

- 3225 • Riya Bhadani
- 3226 • Ameen Agbaria
- 3227 • Jaden Hamer
- 3228 • Pranay Maddireddy

3229 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 3230 • Stance Agreement: 52.4%
- 3231 • Time Agreement: 81.9%
- 3232 • Certainty Agreement: 88%

3233 Annotation Guide

3234 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
3235 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
3236 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
3237 other than monetary policy stance.

3238 The annotation guide for monetary policy stance is summarized in Table 88. It is built by dividing
3239 each target sentence into 12 defined categories: exchange rate, inflation, real GDP, foreign investment,
3240 employment, foreign debt, interest rates, poverty, tourism, oil and natural gas, remittance payments,
3241 and Suez Canal. Each sentence is labeled as hawkish, dovish, neutral, or irrelevant.

- 3242 • *Exchange rate:* Sentence pertaining to the value of Egyptian Pound (EGP) against United
3243 States Dollar (USD).
- 3244 • *Inflation:* A sentence pertaining past or current inflation levels, changes in inflation, or target
3245 inflation.
- 3246 • *Real GDP:* A sentence about changes to Egypt’s economic output or target GDP.

- 3247 • *Employment*: A sentence about changes to the employment levels in Egypt or target employ-
3248 ment levels.
- 3249 • *Foreign Investment*: A sentence about the level of foreign assets Egypt has or changes to
3250 foreign assets.
- 3251 • *Foreign Debt*: A sentence regarding the level of foreign debt in Egypt or changes to foreign
3252 debt.
- 3253 • *Interest Rates*: A sentence pertaining to past or current interest rate levels, changes to interest
3254 rate or target interest rate.
- 3255 • *Poverty*: A sentence pertaining to current poverty levels or changes to poverty levels.
- 3256 • *Tourism*: A sentence pertaining to current tourism revenue or foreseeable or past changes to
3257 tourism revenue.
- 3258 • *Oil and Natural Gas*: A sentence pertaining to changes in oil or natural gas prices, supply,
3259 demand, or exports, or a sentence regarding the affect on Egypt's economy due to global
3260 oil prices or global oil demand. Note oil and natural gas includes crude oil, petroleum, and
3261 petroleum derivatives.
- 3262 • *Remittance Payments*: A sentence pertaining to changes in level of remittance payments as a
3263 function of GDP.
- 3264 • *Suez Canal*: A sentence regarding the number of or change to the number of cargo ships
3265 passing through the Suez Canal, or geopolitical factors that may affect Suez Canal revenue.

3266 **Examples:**

- 3267 • “With the gradual easing of previous shocks, inflationary pressures continued to subside, as
3268 annual headline and core inflation edged downward for the fifth consecutive month to reach
3269 25.7 percent and 24.4 percent in July 2024, respectively.”
3270 **Dovish**: The inflation rate has consistently been decreasing suggesting that the MPC will
3271 decrease interest rates.
- 3272 • “The slight increase in the remittances of the Egyptians working abroad by 1.6 percent to
3273 register US \$ 31.9 billion.”
3274 **Hawkish**: Increase in remittance payments means larger money supply in Egypt and more
3275 economic activity, increase the money supply in the market. Therefore, MPC would increase
3276 interest rates.
- 3277 • “Globally, the expansion of economic activity continued to weaken, financial conditions
3278 eased, and trade tensions continued to weigh on the outlook.”
3279 **Neutral**: Both hawkish and dovish sentiment are expressed. Weakened economic growth
3280 and trade tensions indicates decreasing interest rates but easing financial conditions suggests
3281 that may not be necessary.
- 3282 • “The MPC is watching these developments and is weighing the possible effects of the above
3283 on inflation in the near future.”
3284 **Irrelevant**: There is missing information that makes the Stance unclear.

Table 88: Central Bank of Egypt Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Exchange rate	When exchange rate of EGP against USD increases.	When exchange rate of EGP against USD decreases.	When exchange rate remains stable.	Sentence is not relevant to monetary policy.
Inflation	When inflation or CPI increases or when mention of maintaining a flexible inflation target, "price stability over the medium term", mandate, structural reform, or some adjacent topic.	When inflation or CPI decreases.	When inflation remains unchanged or changes minimally.	Sentence is not relevant to monetary policy.
Real GDP	When real GDP growth is strong.	When real GDP growth rate in Egypt is sluggish.	When GDP remains steady, stable, or unchanged.	Sentence is not relevant to monetary policy.
Employment	When unemployment levels decrease.	When unemployment levels increase.	No or negligible changes to unemployment levels.	Sentence is not relevant to monetary policy.
Foreign Investment	When there is an increase in foreign assets in Egypt.	When there is a decrease in foreign assets in Egypt.	When amount of foreign assets remain unchanged.	Sentence is not relevant to monetary policy.
Foreign Debt	When Egypt reduces its foreign debt or secures debt relief, improving fiscal balance.	When Egypt's foreign debt increases, raising repayment pressure and fiscal instability.	When foreign debt levels remain manageable with regular servicing or foreign debt levels remain unchanged.	Sentence is not relevant to monetary policy.
Interest Rates	When interest rates increase.	When interest rates decrease.	When interest rates remain unchanged.	Sentence is not relevant to monetary policy.
Poverty	When poverty rate decreases indicating economic recovery or raising inflation.	While inflation might stabilize, economic growth slows, increasing poverty levels.	When poverty rate remains unchanged.	Sentence is not relevant to monetary policy.
Tourism	When tourism revenue increases.	When tourism revenue declines.	When tourism revenue remains steady.	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
Oil and Natural Gas	When global oil/gas prices surge, leading to higher energy export revenues but potentially higher domestic energy costs.	When global oil/gas prices drop, lowering export revenues but reducing domestic fuel costs.	When oil/gas prices and exports remain stable, providing steady revenue without major shocks.	Sentence is not relevant to monetary policy.
Remittance payments	When remittance payments decrease as a percentage of GDP.	When remittance payments increase as a percentage of GDP.	When remittance payments remain unchanged.	Sentence is not relevant to monetary policy.
Suez Canal	When the number of cargo ships increases, boosting Suez Canal revenue.	When the number of cargo ships declines, affecting Suez Canal revenue, usually due to a war in the region or blockage in the Red Sea.	When the number of cargo ships passing through the canal doesn't change.	Sentence is not relevant to monetary policy.

3285 The annotation guide for whether the sentence has Forward Looking information or not is summa-
3286 rized in the Table 89 and whether the committee is talking about the sentence with Certainty attest
3287 to it or not in Table 90.

Table 89: Central Bank of Egypt Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	When the sentence discusses future events, economic outlook, possible changes to policy positions, and may use keywords such as “forecast,” “expected,” “judges,” or “future.”	“The Committee reiterates that the path of future policy rates remains a function of inflation expectations.”
Not Forward Looking	When the sentence discusses a past or present economic event, data, or policy position, and may use keywords such as “driven by,” “reflections,” “decided,” “stabilized.”	“Labor market data show that the unemployment rate narrowed to 12.5 percent in the quarter ending June 2016 after peaking at 13.4 percent in the quarter ending December 2013, supported by lower real unit labor costs relative to productivity.”

Table 90: Central Bank of Egypt Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	When the sentence expresses certainty about a future event or outcome, it may use words such as “declined,” “increased,” “decided,” “revised,” or “data show.”	“The overnight deposit and the overnight lending rates were kept at 9.5 percent and 12.5 percent, respectively.”
Uncertain	When the statement expresses ambiguity about future events or causes of a past event, it may use words such as “expected,” “uncertain,” “implies,” “projected.”	“However, economic activity is expected to grow at a slower rate than previously projected, given the uncertainty and negative spillovers from the global scene.”

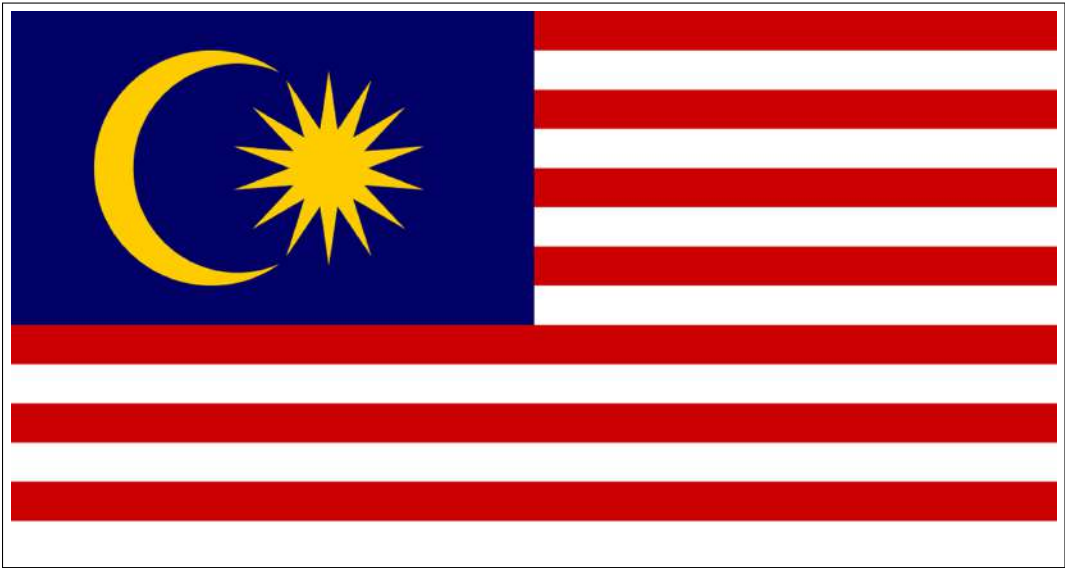
3288

Bank Negara Malaysia

3289

Region: Malaysia

3290



3291

3292

Data Collected: 2004-2024

3293

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

3294

3295 **Monetary Policy Mandate**

3296 The Monetary Policy Committee (MPC) of Bank Negara Malaysia focuses on price stability and
3297 economic growth through a combination of low interest rates and inclusive financial sector.
3298

3299 **Mandate Objectives:**

- 3300 • **Price Stability:** Price stability is achieved by maintaining the Overnight Policy Rate (OPR)
3301 below 3%. OPR influences the interest rates of other areas such as consumer and business
3302 loans.
- 3303 • **Economic Growth:** Economic growth is supported through financial inclusion and providing
3304 ample credit. The bank makes sure every financial sectors have access to financial services
3305 for financial inclusion.

3306 **Structure**

3307 **Composition:**

- 3308 • **Monetary Policy Committee (MPC):** The MPC regulates the monetary policies for the
3309 Bank Negara Malaysia. The MPC is a small group which includes the Governor, Deputy
3310 Governors, and three to seven other members. These members include external members
3311 selected by the Minister of Finance.
- 3312 • **Financial Stability Executive Committee (FSEC):** The FSEC maintains financial stability
3313 by overseeing and deciding on policies related to financial stability. The committee includes
3314 seven members that are mostly independent from the Bank.

3315 **Meeting Structure:**

- 3316 • **Frequency:** The MPC meets a minimum of six times a year.
- 3317 • **Additional Meetings:** The Monetary Policy Working Group decides whether there should
3318 be more meetings for the MPC.

3319 **Manual Annotation**

3320 **Annotators**

- 3321 • Eric Kim
- 3322 • Aiden Chiang
- 3323 • Rudra Gopal
- 3324 • Spencer Gosden

3325 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 3326 • Stance Agreement: 45.8%
- 3327 • Time Agreement: 82.8%
- 3328 • Certainty Agreement: 75.9%

3329 **Annotation Guide**

3330 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
3331 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
3332 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
3333 other than monetary policy stance.

3334 The annotation guide for monetary policy stance is summarized in Table 91. It is built by dividing
3335 each target sentence into thirteen defined categories: Economic status, Malaysian ringgit Value
3336 Change, Household loan, Foreign nations, BNM Expectations/Actions/Assets, Money Supply, Labor,
3337 Interest rate gap between Malaysia and United States, COVID, Energy/Housing Prices, and 1997
3338 Asian Financial Crisis.

- 3339 • *Economic Status*: A sentence pertaining to the state of the economy, relating to unemployment and inflation
- 3340
- 3341 • *Malaysian ringgit Change*: A sentence pertaining to changes such as appreciation or depreciation of value of the Malaysian ringgit on the Foreign Exchange Market
- 3342
- 3343 • *Household loan*: A sentence pertaining to changes in public or private debt levels.
- 3344 • *Foreign Nations*: A sentence pertaining to trade relations between Malaysia and a foreign economy. If not discussing Malaysia we label neutral.
- 3345
- 3346 • *BNM Expectations/Actions/Assets*: A sentence that discusses changes in the BNM yields, bond value, reserves, or any other financial asset value.
- 3347
- 3348 • *Money Supply*: A sentence that overtly discusses impact to the money supply or changes in demand.
- 3349
- 3350 • *Labor*: A sentence that relates to changes in labor productivity.
- 3351 • *Interest rate gap between Malaysia and United States*: A sentence pertaining to whether the US interest rate is lower or higher compared to Malaysia's interest rate.
- 3352
- 3353 • *Energy/Housing Prices*: A sentence pertaining to changes in prices of real estate, energy commodities, or energy sector as a whole.
- 3354
- 3355 • *1997 Asian Financial Crisis*: A sentence pertaining to when the financial crisis caused deflationary pressures or capital outflows.
- 3356

3357 **Examples:**

- 3358 • "At the Monetary Policy Committee (MPC) meeting today, Bank Negara Malaysia decided to leave the Overnight Policy Rate (OPR) unchanged at 3.50 percent."
- 3359 **Neutral**: Maintaining existing interest rates indicates no change in monetary policy.
- 3360
- 3361 • "Notwithstanding the stronger aggregate demand and higher global energy prices, the domestic inflation rate, as measured by the Consumer Price Index, remained low at 1.2% in the second quarter. The underlying inflation in the economy is expected to remain low."
- 3362
- 3363 **Dovish**: Low inflation rates indicate room for economic growth and an accommodative monetary policy.
- 3364
- 3365 • "Over the course of the COVID-19 crisis, the OPR was reduced by a cumulative 125 basis points to a historic low of 1.75% to provide support to the economy. The unprecedented conditions that necessitated such actions have since abated. With the domestic growth on a firmer footing, the MPC decided to begin reducing the degree of monetary accommodation. This will be done in a measured and gradual manner, ensuring that monetary policy remains accommodative to support a sustainable economic growth in an environment of price stability."
- 3366
- 3367 **Hawkish**: Reducing monetary accommodation signals a shift towards a tighter economic policy.
- 3368
- 3369 • "At the current OPR level, the monetary policy stance remains supportive of the economy and is consistent with the current assessment of inflation and growth prospects."
- 3370
- 3371 **Hawkish**: A monetary policy stance that supports growth of the economy indicates an tighter approach.
- 3372
- 3373 • "Headline inflation is projected to average between 2.2% - 3.2% in 2022. Given the improvement in economic activity amid lingering cost pressures, underlying inflation, as measured by core inflation, is expected to trend higher to average between 2.0% - 3.0% in 2022."
- 3374
- 3375 **Hawkish**: A rise in inflation and cost pressure indicates a tighter monetary policy stance to maintain price stability.
- 3376
- 3377 • "The MPC will continue to carefully evaluate the global and domestic economic and financial developments and their implications on the overall outlook for inflation and growth of the Malaysian economy."
- 3378
- 3379 **Neutral**: There is no mention of a policy shift since the MPC is solely observing the situation.
- 3380
- 3381
- 3382
- 3383
- 3384
- 3385
- 3386
- 3387
- 3388
- 3389

Table 91: Bank Negara Malaysia Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Economic Status	When inflation drops below BNM's 1% target, unemployment increases, or economic growth is projected as low.	When inflation exceeds BNM's 3% target, unemployment decreases, economic growth is projected as high, economic output surpasses potential supply, or economic slack declines.	When unemployment rate or growth is unchanged, maintained, or sustained.	Sentence is not relevant to monetary policy.
Malaysian Ringgit Value Change	When the ringgit appreciates beyond BNM targets and inflation is low (below 1.5%), signaling a need to boost growth.	When the ringgit depreciates beyond BNM targets and inflation is high (above 3%), indicating a need for economic cooling.	When inflation is between 1.5% and 3%, and the ringgit's value changes at a healthy rate.	Sentence is not relevant to monetary policy.
Household Loan	When public or private debt levels rise, leading to stricter fiscal policies to manage borrowing (household debt-to-GDP ratio exceeds 84%).	When public or private debt levels decrease, allowing for more borrowing flexibility and fiscal stimulus (household debt-to-GDP ratio drops below 84%).	N/A	Sentence is not relevant to monetary policy.
Foreign Nations	When Malaysia's trade deficit increases with the U.S., Singapore, or Japan, signaling more exports than imports and ringgit appreciation.	When Malaysia's trade deficit rises with China or falls with the U.S., Singapore, or Japan, indicating ringgit depreciation.	When referring to foreign nations' economic or trade policies.	Sentence is not relevant to monetary policy.
BNM Expectations, Actions, and Assets	When BNM expects subpar inflation, predicts disinflation, observes narrowing treasury bond spreads, declining treasury yields, or reduces bank reserves.	When BNM signals tightening, raises interest rates, or increases reserves to curb inflation.	When BNM believes current economic conditions are satisfactory or will self-correct.	Sentence is not relevant to monetary policy.

Category	Dovish	Hawkish	Neutral	Irrelevant
Money Supply	When money supply is low, M2 growth rate falls below 4%, and loan demand increases.	When money supply is high, M2 growth rate exceeds 6%, demand for goods rises, and loan demand is low.	N/A	Sentence is not relevant to monetary policy.
Labor	When productivity increases, solid labor market, strong semiconductor, electronics, and tourism industries, growing wages and consumer spending.	When productivity decreases, tourism is down, exports decreasing, downturn in global technology sector.	N/A	Sentence is not relevant to monetary policy.
Interest rate gap between Malaysia and United States	When US interest rate is aligned with Malaysia's, US not interfering with Malaysia's interest rates.	When US interest rate is higher than Malaysia, potentially pressuring Malaysia to raise interest rates.	N/A	Sentence is not relevant to monetary policy.
COVID	When easing Covid regulation: boost tourism industry and increased productivity.	When increased Covid: regulation detrimental for tourism industry and decreased productivity.	N/A	Sentence is not relevant to monetary policy.
Energy and Housing Prices	When oil/energy prices decrease, when house prices decrease or expected to decrease.	When oil/energy prices increase, when house prices increase or expected to increase.	N/A	Sentence is not relevant to monetary policy.
1997 Asian Financial Crisis	When the financial crisis causes deflationary pressures, prompting the central bank to lower interest rates to stimulate demand and growth.	When the financial crisis causes capital outflows, and the central bank raises interest rates to defend the currency and attract foreign capital.	N/A	Sentence is not relevant to monetary policy.

3390 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 92 and whether the committee is talking about the sentence with Certainty attest
3391 to it or not in Table 93.
3392

Table 92: Bank Negara Malaysia Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	When a sentence contains phrases such as “expected,” “anticipated,” “projected,” “will,” or “forecasted.”	“Budget 2025 measures will provide additional support to growth.”
Not Forward Looking	When a sentence contains phrases such as “had been,” “previously,” or “last year.”	“Domestic headline inflation moderated in September to 1.8.”

Table 93: Bank Negara Malaysia Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	When a sentence contains phrases such as “will,” “must,” “certainly,” or “definitely.”	“The monetary policy statement will be released at 6 p.m. on the same day as the MPC meeting.”
Uncertain	When a sentence contains phrases such as “might,” “could,” “may,” “possibly,” “likely,” or “uncertain.”	“These developments may increase the risk to the outlook for inflation.”

3393
3394
3395

3396
3397

Central Bank of the Philippines

Region: Philippines



Data Collected: 2002-2024

3398
3399

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

3400 **Monetary Policy Mandate**

3401 **Mandate Objectives:** The BSP's primary objective is to promote price stability that fosters balanced
3402 and sustainable growth of the economy.

3403 **Structure**

3404 **Composition:** The BSP's Monetary Board contains seven members, who are all appointed by the
3405 President of the Philippines. The seven members are:

- 3406 • The Governor of the BSP, who also serves as the Chairman of the Monetary Board for a
3407 six-year term.
- 3408 • A member of the Cabinet who serves a six-year term.
- 3409 • Five full-time members, three of whom serve a six-year term, and the other three serve a
3410 three-year term.

3411 **Meeting Structure:**

- 3412 • **Frequency:** Six scheduled meetings annually (bi-monthly). The number of meetings
3413 decreased from seven starting in 2025.
- 3414 • **Additional Meetings:** The Governor of the BSP or two other board members may call a
3415 meeting whenever necessary to address urgent economic situations.

3416 **Manual Annotation**

3417 **Annotators**

- 3418 • Vedant Nahar
- 3419 • Saloni Jain
- 3420 • Saketh Budideti
- 3421 • Jiashen Zhang

3422 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 3423 • Stance Agreement: 53.5%
- 3424 • Time Agreement: 88.1%
- 3425 • Certainty Agreement: 80.7%

3426 **Annotation Guide**

3427 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
3428 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
3429 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
3430 other than monetary policy stance.

3431 The annotation guide for monetary policy stance is summarized in Table 94. It is built by dividing
3432 each target sentence into five defined categories: Policy, Environment, Society, Market, and Trade.

- 3433 • *Policy:* Decisions implemented by the government or its institutions to regulate, direct, or
3434 influence national affairs. Policies often impact laws and public services.
- 3435 • *Environment:* Decisions related to environmental sustainability, energy consumption, or
3436 natural resources.
- 3437 • *Society:* Decisions aimed at addressing the well-being of society. Policies often address
3438 income inequality, cost-of-living, and social programs.
- 3439 • *Market:* Decisions that influence financial markets, including changes in interest rates, bond
3440 yields, and liquidity management.

3441 • *Trade*: Decisions that impact trade, including imports, exports, exchange rates, and trade
3442 balances.

3443 **Examples:**

3444 • “Maintain the BSP’s policy interest rates at 7.5 percent for the overnight RRP (borrowing)
3445 rate and 9.75 percent for the overnight RP (lending) rate; b.”

3446 **Neutral**: Maintaining existing interest rates indicates no change in monetary policy.

3447 • “The Monetary Board (MB) decided to: a) Increase the BSP’s key policy interest rate by
3448 50 basis points to 4.50 percent for the overnight RRP (borrowing) facility, effective 28
3449 September 2018; and, b) Adjust the interest rates on the overnight deposit and overnight
3450 lending facilities accordingly.”

3451 **Hawkish**: Increase in interest rate and Reverse Repurchase Rate (RRP) signals contrac-
3452 tionary monetary policy.

3453 • “In August, the peso depreciated against the US dollar amid renewed global recession
3454 concerns as a result of protracted trade conflict between the US and China.”

3455 **Hawkish**: A depreciating peso in addition to recession concerns is likely to drive inflation
3456 higher, creating a need for contractionary monetary policy to ensure the stability of the peso
3457 and the economy.

3458 • “Moreover, unemployment rose to 10.2 percent in the fourth quarter of 2002 as against 9.8
3459 percent in the same period a year ago.”

3460 **Dovish**: A sharp increase in the unemployment rate signals monetary easing to stimulate
3461 the economy.

3462 • “Compared to the previous policy meeting, baseline inflation forecast for 2013 has been
3463 revised downward due primarily to lower oil prices.”

3464 **Dovish**: Lower inflation forecasts indicate room for monetary easing without the risks of
3465 high inflation.

3466 • “The highlights of the discussions during the 9 November 2017 monetary policy meeting
3467 were approved by the Monetary Board during its regular meeting held on 28 November
3468 2017.”

3469 **Irrelevant**: The sentence mentions the approval of a past discussion with no mention of a
3470 monetary policy stance.

Table 94: Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Policy	When government spending decreases or financial regulations are tightened.	When government spending increases or financial regulations are loosened.	When policy statements maintain current spending levels or regulations.	Sentence is not relevant to monetary policy.
Environment	Energy prices fall, new price controls are implemented on fossil fuels, added subsidies for renewable energy, adverse effects of el niño.	Energy prices rise, production costs rise, or subsidies are removed.	When statements maintain current prices, confirm existing policy (e.g. subsidies, price controls) without signaling change.	Sentence is not relevant to monetary policy.
Society	When there is higher taxes, reduced subsidies, reduced social programs, or reduced minimum wages.	When there are lower taxes, increased subsidies, new social programs, or increases in minimum wages.	When societal measures are discussed with no clear indication of change to fields such as taxes, minimum wage, subsidies, or social programs.	Sentence is not relevant to monetary policy.
Market	When there is a reduction in interest rates, increase in the money supply, higher demand for loans, narrowing bond spread, higher bond yields.	When there is an increase in interest rates, decrease in money supply, lower demand for loans, expanding bond spread, lower bond yields.	When there is no change in interest rates, money supply, or demand for loans.	Sentence is not relevant to monetary policy.
Trade	Philippine peso appreciates, increased trade deficit, decreased exports.	Philippine peso depreciates, increased trade surplus, increased imports.	Maintained trade activity and currency valuation.	Sentence is not relevant to monetary policy.

3471 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 95 and whether the committee is talking about the sentence with Certainty attest
3472 to it or not in Table 96.
3473

Table 95: Bangko Sentral ng Pilipinas Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	text describing events anticipated to occur in the future. Sentences often utilize language such as “expected” or “projected.”	“These developments form the economic backdrop for the latest WEO projections by the IMF, which point to a modest recovery in 2017.”
Not Forward Looking	text referring to past or recent economic policy, data, trends, or events.	“Compared to the survey two quarters ago, fewer banks have tightened lending standards over the past six months based on the Fed’s Senior Loan Officer Survey.”

Table 96: Bangko Sentral ng Pilipinas Annotation Guide to identify whether a sentence echoes certainty or not.

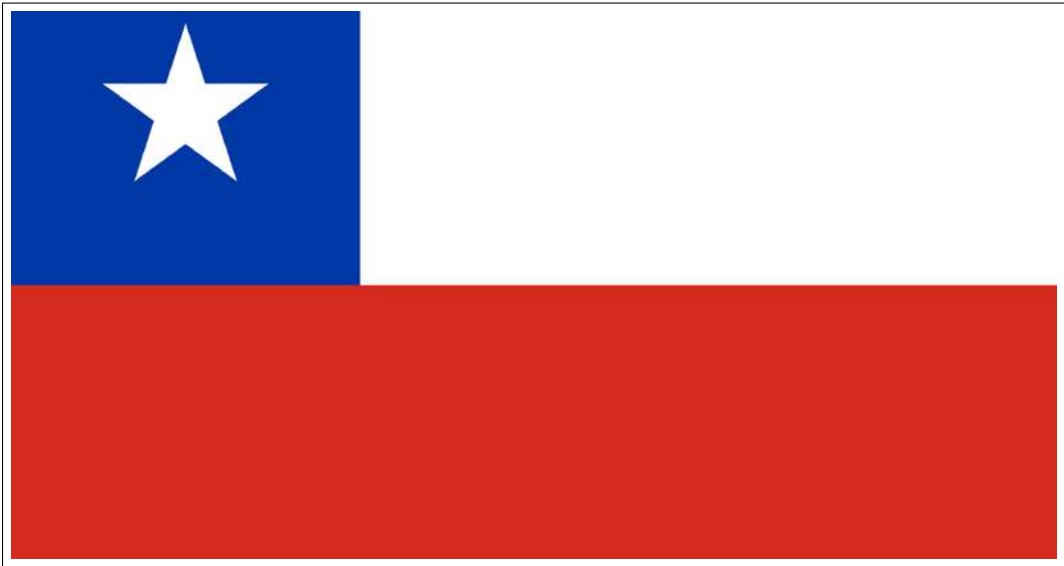
Label	Description	Example
Certain	The sentence reflects a clear, firm, or definitive decision or expectation about actions or outcomes without ambiguity.	“Corporate bond issuances also increased to P119 billion, higher by 20 percent compared to the level a year ago.”
Uncertain	The sentence expresses potential risks, varying opinions, or uncertainties about events, outcomes, or actions, often using words like “could,” “might,” or “may.”	“However, a number of risks remain, notably the expected slower growth trajectory of the US economy, a disorderly unwinding of global imbalances, and a potential increase in financial market volatility.”

3474
3475
3476

3477
3478

Central Bank of Chile

Region: Chile



Data Collected: 2018-2024

3479
3480

Important Links
[Central Bank Website](#)
[Annotated Dataset](#)
[Stance Label Model](#)
[Time Label Model](#)
[Certain Label Model](#)

3481 Monetary Policy Mandate

3482 The primary mandate of the Bank of Chile is to control the inflation levels in the economy.

3483 Mandate Objectives:

- 3484 • **Price Stability:** The Central Bank of Chile's primary goal is to ensure stability of the
3485 currency. This implies keeping inflation low and stable over time. This translates to keeping
3486 the Consumer Price Inflation at around 3% with a $\pm 1\%$ tolerance over a two year horizon.

3487 Structure

3488 Composition:

- 3489 • **Board of the Central Bank of Chile:** The Board is made up of five members, appointed by
3490 the President of the Republic, upon prior approval by the Senate. Each member serves for
3491 ten years. At the same time, one of them serves as the Governor for a term of five years.
- 3492 • **Additional Members:** Bank managers and analysts also participate in the meetings, along
3493 with a representative of the Ministry of Finance, all of them without voting rights.

3494 Meeting Structure:

- 3495 • **Frequency and Length:** There are 8 meetings that occur every year. Four of these meetings
3496 are held on the business day prior to the publication of the Monetary Policy Report and they
3497 take place on one day. On the other four occasions, the meeting lasts one and a half days.

3498 Manual Annotation

3499 Annotators

- 3500 • Sagnik Nandi
- 3501 • Rutwik Routu
- 3502 • Ankith Keshireddy
- 3503 • Dylan Patrick Kelly

3504 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 3505 • Stance Agreement: 55.5%
- 3506 • Time Agreement: 69.6%
- 3507 • Certainty Agreement: 75.5%

3508 Annotation Guide

3509 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
3510 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
3511 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
3512 other than monetary policy stance.

3513 The annotation guide for monetary policy stance is summarized in Table 97. It is built by dividing
3514 each target sentence into nine defined categories: Inflation, Economic Growth, Copper Prices, Peso
3515 Exchange Rate, Foreign Reserves, External Demand (Trade Partners), Key Words/Phrases, Labor
3516 Market, Energy Prices.

- 3517 • *Inflation:* A sentence relating to the monetary policy decision of the Central Bank of Chile.
- 3518 • *Economic Growth:* A sentence pertaining to the state of the economy, relating to unemploy-
3519 ment, investment, and inflation.
- 3520 • *Copper Prices:* A sentence relating to changes in copper prices, since Chile is the largest
3521 copper exporter and therefore, is an economic indicator for Chile.

- 3522 • *Peso Exchange Rate*: A sentence pertaining to changes in the peso, such as appreciation or
3523 depreciation of the value of the Peso on the Foreign Exchange Market.
- 3524 • *Foreign Reserves*: A sentence reflecting the central bank's stance on liquidity and currency
3525 stabilization.
- 3526 • *External Demand*: Since Chile's economy is export-dependent, a sentence pertaining to
3527 global demand becomes a critical factor.
- 3528 • *Key Words/Phrases*: A sentence that contains a key word or phrase that would classify it
3529 squarely into one of the three label classes, based upon its frequent usage and meaning
3530 among particular label classes.
- 3531 • *Labor Market*: A sentence that relates to changes in labor productivity.
- 3532 • *Energy Prices*: Lower energy prices reduce inflation and production costs, while higher
3533 energy prices increase them, prompting different monetary responses from the Central Bank.

3534 **Examples:**

- 3535 • "Meanwhile, the longer-term rates had risen less than their external counterparts, in a context
3536 in which the local risk indicators remained contained."
3537 **Neutral**: This statement only describes the market conditions without indicating a tightening
3538 or easing policy stance.
- 3539 • "The median of the Financial Traders Survey (FTS) remained around 3.5% for the third
3540 month in a row, while 90% of the respondents expected inflation to be higher than 3%."
3541 **Hawkish**: The fact that 90% of the respondents believed that the inflation was above 3%
3542 suggests that there are some inflationary pressures.
- 3543 • "In this context, a major challenge would be to maintain the impulse of monetary and fiscal
3544 policy, as well as the ability to safeguard financial stability, until the economy was able to
3545 achieve a self-sustained growth rate, reducing the gaps and frictions that still remained."
3546 **Dovish**: The statement suggests a goal to maintain stability which suggests a more accom-
3547 modative approach.
- 3548 • "All five board members agreed that, from the analysis of the background information
3549 submitted in the preparation of the March MP report, it could be concluded that there was
3550 still no evidence of consolidation of said inflationary convergence."
3551 **Hawkish**: In this sentence, it was evident from the board members' agreement that the
3552 inflation has not yet stabilized, signaling that inflation is too high or volatile. Thus, potentially
3553 requiring tightening monetary policy.
- 3554 • "There were also differences across multiple sources of information."
3555 **Irrelevant**: This sentence points out that differences exist across information sources but is
3556 ultimately carry any monetary policy stance.

Table 97: Central Bank of Chile Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Inflation	Sentences indicating inflation significantly below the 2% target suggest concerns over deflation or weak demand.	Sentences where inflation exceeds 4% highlight the necessity for tighter monetary policies to curb inflationary pressures.	Sentences indicating inflation rates within the 2.5% to 3.5% ranges suggest stable price levels with no immediate need for policy adjustments.	Sentence is not relevant to monetary policy.
Economic Growth	Descriptions of GDP growth below 2% or unemployment over 8% indicate the need for economic stimulus.	Descriptions of GDP growth above 5% and unemployment below 4% suggest an overheating economy requiring cooling measures.	Sentences indicating GDP growth between 3% and 4%, and unemployment rates from 4% to 6%, represent balanced economic conditions.	Sentence is not relevant to monetary policy.
Copper Prices	Statements that copper prices are below approximately \$7,000/ton indicate reduced export revenues and potential economic strain.	References to copper prices exceeding \$10,000/ton suggest potential inflationary pressures, possibly necessitating tighter monetary policies.	Descriptions of copper prices within \$7,500 to \$9,500/ton indicate stable conditions conducive to current policy settings.	Sentence is not relevant to monetary policy.
Peso Exchange Rate	Any mention of the peso depreciating more than 5% suggests potential harm to export competitiveness, which might call for central bank intervention.	Mentions of the peso depreciating more than 5%-10% in a short period, pointing to possible inflation spikes and the need for a tighter monetary stance.	Statements indicating minor fluctuations (less than 2% per month) suggest a stable exchange rate.	Sentence is not relevant to monetary policy.
Foreign Reserves	Rapid increases in reserves (more than 5% per month) can indicate an intention to inject liquidity into the economy.	Decreases in reserves (more than 5%) can suggest a tightening of monetary liquidity.	Moderate changes (within 2% month-over-month) in reserves indicate stability.	Sentence is not relevant to monetary policy.
External Demand (Trade Partners)	Sentences indicating a weakening in external demand, especially from major trade partners like China, suggest economic downturns.	Increased global demand for Chilean exports, as mentioned in statements, indicates robust economic health.	Stable external demand levels imply a steady economic environment.	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
Key Words and Phrases	Sentences focusing on “stimulating growth,” “accommodative monetary policy,” or “reducing unemployment” are categorized as dovish.	Phrases like “inflation control,” “price stability,” or “reducing liquidity” are considered hawkish.	Neutral terms include “balanced,” “steady,” or “in line with expectations.”	Sentence is not relevant to monetary policy.
Labor Market	References to high unemployment rates (over 8%) and stagnant wage growth (less than 2%) indicate a need for supportive monetary policies.	Low unemployment (less than 4%) and rising wages (over 5%) suggest an overheating labor market that may need cooling measures.	Descriptions of unemployment rates between 4%-6% and wage growth between 3%-4% indicate a healthy labor market.	Sentence is not relevant to monetary policy.
Energy Prices	Statements of energy prices declining by more than 10% often suggest decreasing inflationary pressures and reduced production costs.	Descriptions of energy prices increasing by more than 10% indicate rising inflation pressures, potentially requiring tighter monetary measures.	Sentences noting energy price changes within a 1%-2% range month-over-month suggest price stability.	Sentence is not relevant to monetary policy.

3557 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 98 and whether the committee is talking about the sentence with Certainty attest
3558 to it or not in Table 99.
3559

Table 98: Central Bank of Chile Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Sentences that contain predictions or anticipations about future events or trends.	“However, the latest indicators showed a rebound in confidence and outlook, anticipating a faster recovery in the second half of the year, supported by progress in the vaccination processes, the high savings accumulated by households, and the fiscal stimulus plans.”
Not Forward-looking	Sentences that discuss current or past events without reference to future expectations.	“Household expectations had also improved, although their performance and that of other indicators more closely related to private consumption augured a slow rebound in this expenditure component.”

Table 99: Central Bank of Chile Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Sentences that state facts or outcomes with certainty, without speculative language.	“Headline inflation had reached 13.7% annually in September, slightly less than in August, while core inflation – the CPI minus volatiles – had risen to 11.1% annually.”
Uncertain	Sentences that include words indicating probability, speculation, or uncertainty such as “likely,” “could,” or “might.”	“It was also noted that the external risks had implications for inflation and monetary policy in Chile that could differ in both the short and the long term.”

3560

Central Reserve Bank of Peru

3561

Region: Peru



3563

3564

Data Collected: 2001-2024

- 3565
- 3566
- Important Links**
- [Central Bank Website](#)
 - [Annotated Dataset](#)
 - [Stance Label Model](#)
 - [Time Label Model](#)
 - [Certain Label Model](#)

3567 **Monetary Policy Mandate**

3568 The BCRP (Banco Central de Reserva del Peru) is tasked with regulating the amount of money,
3569 manage international reserves, issue notes and coins, and periodically report on the economy's
3570 finances

3571 **Mandate Objectives:**

- 3572 • **Price Stability:** Maintain price stability via inflation targeting 2% per year with $\pm 1\%$.
- 3573 • **Interbank Interest Rate:** Either injecting or sterilizing liquidity depending on whether the
3574 reference rate is facing upward or downward pressure.

3575 **Structure**

3576 BCRP's was established by the constitution of Peru and serves to maintain monetary stability in the
3577 economy.

3578 **Composition:**

- 3579 • **The Board of Directors:** They have the highest authority of the central bank. The board is
3580 responsible for determining the policies needed to fulfill the bank's mission and also is in
3581 charge of general activities. The Executive and Legislative branches of Peru appoint three
3582 members each to the board.
- 3583 • **General Manager:** The General Manager is in charge of the technical and administrative
3584 processes.

3585 **Meeting Structure:**

- 3586 • **Frequency:** The Board of Directors meet once a month or 12 times a year at dates published
3587 at the start of the year.
- 3588 • **Purpose:** The meetings general discuss what actions need to be taken in order to meet the
3589 bank goals. The decisions made during the meeting are immediately announced to the press
3590 afterward.
- 3591 • **Method:** Monetary policy actions that can be taken include changing the reference rate
3592 depending on inflationary or deflationary pressures.

3593 **Manual Annotation**

3594 **Annotators**

- 3595 • Meghaj Tarte
- 3596 • Min Lee
- 3597 • Saketh Budideti
- 3598 • Sloka Chava

3599 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 3600 • Stance Agreement: 56.6%
- 3601 • Time Agreement: 82.2%
- 3602 • Certainty Agreement: 87.6%

3603 **Annotation Guide**

3604 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
3605 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
3606 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
3607 other than monetary policy stance.

3608 The annotation guide for monetary policy stance is summarized in Table 100. It is built by dividing
 3609 each target sentence into twelve defined categories: Economic Status, Sol Value Change, Housing
 3610 Debt/Loan, Foreign Nations, BCRP Expectations/Actions/Assets, Money Supply, Keywords/Phrases,
 3611 Dollar Value, FX Market, COVID, Energy/Housing Prices, and Weather.

- 3612 • *Economic Status*: A sentence pertaining to the state of the economy, relating to unemploy-
 3613 ment and inflation.
- 3614 • *Sol Value Change*: A sentence pertaining to changes such as appreciation or depreciation of
 3615 value of the Peruvian sol on the Foreign Exchange Market
- 3616 • *Housing Debt/Loan*: A sentence that relates to changes in debt in households.
- 3617 • *Foreign Nations*: A sentence pertaining to trade relations between Peru and a foreign
 3618 economy. If not discussing Peru we label neutral.
- 3619 • *BCRP Expectations/Actions/Assets*: A sentence that discusses changes in CRBP yields, bond
 3620 value, reserves, or any other financial asset value.
- 3621 • *Money Supply*: A sentence that overtly discusses impact to the money supply or changes in
 3622 demand.
- 3623 • *Keyword/Phrases*: A sentence that contains key word or phrase that would classify it
 3624 squarely into one of the three label classes, based upon its frequent usage and meaning
 3625 among particular label classes
- 3626 • *Dollar value*: A sentence about fluctuation of the dollar's value which can influence monetary
 3627 policy in a dovish or hawkish direction.
- 3628 • *Foreign Exchange (FX) Market*: A sentence about stability's changes in the foreign exchange
 3629 market
- 3630 • *COVID*: A sentence indicating whether the government is tightening or easing COVID
 3631 policy
- 3632 • *Energy/Housing prices*: A sentence pertaining to changes in prices of real estate, energy
 3633 commodities, or energy sector as a whole.
- 3634 • *Weather*: A sentence pertaining to how drastic changes to weather affect the economy.

3635 Examples:

- 3636 • “In light of the narrowing interest rate gap between Peru and the United States, some mem-
 3637 bers suggested that the Central Reserve Bank of Peru may consider keeping interest rates
 3638 low to support domestic economic recovery and encourage borrowing.”
 3639 **Dovish**: This statement supports economic recovery via lowering interest rates, thus, reflect-
 3640 ing a dovish stance.
- 3641 • “This measure is mainly preventive in a context of strong dynamism of domestic demand, a
 3642 situation in which withdrawing monetary stimulus is advisable in order to maintain inflation
 3643 within the target range.”
 3644 **Hawkish**: This sentence suggests 'withdrawing monetary stimulus', which implies tighten-
 3645 ing monetary policy.
- 3646 • “Therefore, between May 2022 and March 2024, the minimum legal reserve requirement
 3647 rate in local currency was kept at 6 percent in the context of a progressive reduction of the
 3648 monetary easing linked to the COVID-19 epidemic. This rate was lowered to 5.5 percent
 3649 in April of this year in order to increase the amount of money that may be lent out and to
 3650 supplement monetary easing.”
 3651 **Dovish**: This statement indicates a dovish stance as it suggests monetary easing and lowering
 3652 of the legal reserve requirement rate to increase the money available for lending.
- 3653 • “This higher level is explained by the new methodology that is being used to calculate this
 3654 indicator since September 29, and now includes bonds involving a longer maturity in the
 3655 basket”
 3656 **Neutral**: This statement is purely informational, describing the change in how an indicator
 3657 is calculated without any discussion of policy implications.

3659
3660
3661
3662

- “The board will approve the monetary program for the next month on its session of November 5.”
Irrelevant: This statement is purely administrative, providing details about a scheduled board session without indicating any monetary policy stance.

Table 100: Central Reserve Bank of Peru Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Economic Status	When inflation decreases, when unemployment increases, or when economic growth is projected as low.	When inflation increases, unemployment decreases, or economic growth is projected as high.	When unemployment rate or growth is unchanged, maintained, or sustained.	Sentence is not relevant to monetary policy.
Sol Value Change	When Sol value appreciates.	When Sol value depreciates.	N/A	Sentence is not relevant to monetary policy.
Housing Debt/Loan	When public or private debt levels decrease, signaling more flexibility for borrowing and fiscal stimulus.	When public or private debt levels increase, prompting strict fiscal policy to limit borrowing and manage rising debt.	N/A	Sentence is not relevant to monetary policy.
Foreign Nations	When the Peruvian trade deficit decreases.	When the Peruvian trade deficit increases.	When relating to a foreign nation's economic or trade policy.	Sentence is not relevant to monetary policy.
BCRP Expectations, Actions, and Assets	When BCRP expects subpar inflation, disinflation, narrowing spreads of treasury bonds, decreases in treasury security yields, or reductions in bank reserves.	When BCRP signals the need for tightening, raising interest rates, or increasing reserves to curb inflation.	N/A	Sentence is not relevant to monetary policy.
Money Supply	When the money supply is low, M2 increases, or demand for loans increases.	When the money supply is high, demand for goods increases, or demand for loans decreases.	N/A	Sentence is not relevant to monetary policy.
Keywords and Phrases	When the stance is "accommodative," indicating a focus on "maximum employment" and "price stability."	Indicating a focus on "price stability" and "sustained growth."	Use of phrases such as "mixed," "moderate," or "reaffirmed."	Sentence is not relevant to monetary policy.
Dollar Value	When USD is moderate or weakening.	When USD strengthens, indicating potential depreciation of Sol.	N/A	Sentence is not relevant to monetary policy.

Category	Hawkish	Dovish	Neutral	Irrelevant
FX Market	Stable FX market, leading to less government involvement.	Unstable FX market, leading to more government involvement.	N/A	Sentence is not relevant to monetary policy.
COVID	Easing COVID regulations boosts the tourism industry and increases productivity.	Increased COVID regulations harm the tourism industry and decrease productivity.	N/A	Sentence is not relevant to monetary policy.
Energy and Housing Prices	When oil/energy prices decrease or house prices decrease or are expected to decrease.	When oil/energy prices increase or house prices increase or are expected to increase.	N/A	Sentence is not relevant to monetary policy.
Weather	Favorable weather conditions benefit agriculture, fishing, and manufacturing sectors.	Unfavorable weather conditions, such as droughts, harm agriculture, fishing, and manufacturing sectors.	N/A	Sentence is not relevant to monetary policy.

3663 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 101 and whether the committee is talking about the sentence with Certainty
3664 attest to it or not in Table 102.
3665

Table 101: Central Reserve Bank of Peru Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Events that are referencing future actions or future data projections. Sentences using language such as “expected,” “projected,” “coming,” and “next.”	“The BCRP expects to drop interest rates in the next coming meetings.”
Not Forward Looking	Events that are referencing past actions or past data. Sentences using language such as “was,” “stayed,” “maintained,” and “kept.”	“The BCRP decreased the reference rate by 25 bps.”

Table 102: Central Reserve Bank of Peru Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	The sentence is very clear about the stance it is taking or it is reflecting about past actions which are certain as they have already been done. Key words include “certain,” “evidently,” and “absolutely.”	“The BCRP will not change the reference rate at this time.”
Uncertain	The sentence is unsure about the stance it is taking or its speculating about the future which can be hard to predict. Key words include “unsure,” “might,” “may,” “expects,” and “hopes.”	“The BCRP may increase rates at the next meeting.”

3666

3667

3668

3669

3670

Bank of the Republic

Region: Colombia



Data Collected: 2014-2024

3671

3672

Important Links

[Central Bank Website](#)

[Annotated Dataset](#)

[Stance Label Model](#)

[Time Label Model](#)

[Certain Label Model](#)

3673 Monetary Policy Mandate

3674 The Banco de la República (BanRep), in addition to being responsible for maintaining the functioning
3675 of payment systems and managing cultural activities, aims to formulate monetary policy to achieve
3676 its mandate.

3677 Mandate Objectives:

- 3678 • **Maintaining purchasing power** Aims to maintain a low and stable annual inflation target
3679 of 3%, measured by the CPI.
- 3680 • **Economic Growth:** Strives for economic output to be as close to its maximum sustainable
3681 potential value as possible.
- 3682 • **Financial Stability:** Promote the ability of the Colombian economy to allocate resources
3683 and mitigate financial risks properly.

3684 Structure

3685 Composition:

- 3686 • Seven directors, including the Minister of Finance, the Governor of the Bank (who is
3687 elected by the Board), and five members appointed by the President, each serving four-year
3688 renewable terms. Two board members are replaced every presidential term.

3689 Meeting Structure:

- 3690 • **Meeting Requirements:**
 - 3691 – At least five members must be in attendance, one of whom is the the Minister of
3692 Finance.
 - 3693 – Decisions require a majority vote of at least four members, and national loans require a
3694 unanimous vote.
- 3695 • **Frequency:** Monthly for a total of 12 meetings a year. However, BanRep generally does not
3696 make decisions in May, August, or November.

3697 Manual Annotation

3698 Annotators

- 3699 • Joshua Zhang
- 3700 • Saksham Purbey
- 3701 • Adhish D Rajan
- 3702 • Anir Vetukuri

3703 **Annotation Agreement** The agreement percentage among the pairs of annotators for different labels.

- 3704 • Stance Agreement: 57.3%
- 3705 • Time Agreement: 80.4%
- 3706 • Certainty Agreement: 79.9%

3707 Annotation Guide

3708 Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify their annotation guide
3709 further, keeping in mind the changing global macroeconomic dynamics. Also, we have two additional
3710 categories (Forward Looking or Not Forward Looking and Certain or Uncertain) of labels
3711 other than monetary policy stance.

3712 The annotation guide for monetary policy stance is summarized in Table 103. It is built by dividing
3713 each target sentence into eight defined categories: Inflation, Employment, Growth, Peso Value
3714 Change, Trade, BanRep Interest Rate, Energy, and Climate.

- 3715 • *Inflation*: How much inflation is present in the Colombian Economy, measured by CPI, and
3716 whether or not this value increases.
- 3717 • *Employment*: The rate of unemployment in Colombia and its change over time.
- 3718 • *Growth*: The rate of GDP growth in Colombia's economy, and any changes to its value.
- 3719 • *Peso Value Change*: The appreciation, depreciation, or maintenance of the Colombian Peso.
- 3720 • *Trade*: The amount of imports and exports going in, and any changes in the amounts and
3721 distribution of these goods.
- 3722 • *BanRep Interest Rate*: The current interest rate set by the Central Bank of Colombia and any
3723 decisions surrounding it.
- 3724 • *Energy*: The availability and price of energy in Colombia
- 3725 • *Climate*: Notable climate conditions impacting Colombia, such as El Niño.

3726 **Examples:**

- 3727 • "They noted that the concern of a significant impact on food or power costs by the El Niño
3728 phenomenon has been waning."
- 3729 **Dovish**: Poor climate conditions like El Niño put inflationary pressures on agricultural and
3730 energy sectors. With these effects improving, there is less inflation and more room to cut
3731 rates to stimulate the economy.
- 3732 • "They cautioned against sudden and unexpected interest rate cuts, which could fuel inflation
3733 expectations and currency depreciation, potentially jeopardizing the approximation to the
3734 inflation target and necessitating reversals with significant credibility costs."
- 3735 **Hawkish**: The BanRep committee members state their hesitancy toward cutting rates,
3736 indicating a hawkish outlook.
- 3737 • "In the case of electricity rates, they emphasize that the risk of power outages has dissipated
3738 as hydro generation has been increasing and new solar plants are coming online."
- 3739 **Dovish**: Better power infrastructure is less likely to put inflationary pressures on energy
3740 prices, reducing the need to increase rates.
- 3741 • "This trend persists because essential items like rents, education, and personal services,
3742 among others, tend to adjust in line with observed inflation and wage increments."
- 3743 **Neutral**: Makes a neutral remark regarding observed trends surrounding inflation, but does
3744 not provide a reason to either cut or raise rates.
- 3745 • "A regular meeting of the board of directors of Banco de la Republica was held in the city of
3746 Bogota on may 23, 2008."
- 3747 **Irrelevant**: This sentence is a statement about an administrative meeting and does not
3748 contain any monetary policy stance.

Table 103: Bank of the Republic Annotation Guide

Category	Dovish	Hawkish	Neutral	Irrelevant
Inflation	When inflation is declining or is below the 3% target.	When inflation is rising or is above the 3% target.	When inflation is close to the target and unchanged.	Sentence is not relevant to monetary policy.
Employment	Low/decreasing employment, deterioration of job quality, or sectors lobbying for rate cuts to improve the labor market.	High/increasing employment or improvements in the labor market.	Stable unemployment rates.	Sentence is not relevant to monetary policy.
Economic Growth	When economic growth is currently low or expected to be low in the future.	When economic growth is currently too high to be sustainable, or expected to be as such.	When economic growth is stable without causing economic overheating.	Sentence is not relevant to monetary policy.
Peso Value Change	When the Peso appreciates.	When the Peso depreciates.	N/A	Sentence is not relevant to monetary policy.
Trade	More exports or fewer imports, leading to a lower trade deficit.	Less exports or more imports, leading to a higher trade deficit.	N/A	Sentence is not relevant to monetary policy.
BanRep Interest Rate	When BanRep decides to or is planning on lowering the interest rate.	When BanRep decides to or is planning on raising the interest rate.	When BanRep leaves interest rate unchanged.	Sentence is not relevant to monetary policy.
Energy	Falling energy prices, higher energy availability.	Rising energy prices, lower energy availability.	N/A	Sentence is not relevant to monetary policy.
Climate	When climate in the region improves, reducing agricultural risks leading to inflation (for examples, food supply and prices).	When poor climate and weather conditions increase food prices or reduce agricultural output, leading to inflation.	N/A	Sentence is not relevant to monetary policy.
Investment	Lower levels of investment or delays in investment.	Higher levels of investment.	N/A	Sentence is not relevant to monetary policy.
Security Prices	Higher prices of public debt.	Lower prices of public debt.	N/A	Sentence is not relevant to monetary policy.
Credit	Decrease in credit or low levels of credit growth.	When credit growth is high.	N/A	Sentence is not relevant to monetary policy.

3749 The annotation guide for whether the sentence has Forward Looking information or not is sum-
3750 marized in the Table 104 and whether the committee is talking about the sentence with Certainty
3751 attest to it or not in Table 105.

Table 104: Bank of the Republic Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Discusses projections or expectations of future values of metrics. Uses words like “expects,” “forecasts,” or “anticipates.”	“Economic activity indicators suggest that GDP would remain on a recovery path during the second quarter, with varying performances among sectors.”
Not Forward Looking	Discusses historical or ongoing data/events. Uses words like “previous,” “historic” or “recorded.”	“The directors agree that the restrictive monetary policy stance has fulfilled its objective of reducing excess spending in the economy.”

Table 105: Bank of the Republic Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	Refers to concrete decisions, historical facts/data, or projections for the future with definite and unambiguous diction. Uses words like “firmly,” “will,” or “decided.”	“The Board of Directors reached a unanimous decision to lower interest rates to support economic recovery and address inflation, which is effectively decreasing and approaching the target.”
Uncertain	Refers to concrete decisions, historical facts/data, or projections for the future with ambiguous diction and a lack of definite terms. Uses words like “may,” “possible,” or “suggested.”	“They cautioned against sudden and unexpected interest rate cuts, which could fuel inflation expectations and currency depreciation, potentially jeopardizing the approximation to the inflation target and necessitating reversals with significant credibility costs.”

3752

3753



3756

3757

3758

3759 Monetary Policy Mandate

3760 The Czech National Bank (CNB) is responsible for maintaining monetary and financial stability in
3761 the Czech Republic while promoting sound economic development.

3762 Mandate Objectives:

- 3763 • **Inflation Control:** Managing inflation within the CNB's target range, which is currently
3764 announced as 2%.
- 3765 • **Currency management:** Issuing legal tender and managing the Czech money supply.
- 3766 • **Price and Financial Stability:** Ensuring currency stability and the stability of the broader
3767 Czech financial system through effective monetary policy decisions.

3768 Structure

3769 Composition:

- 3770 • **Governor:** Represents the Czech National Bank externally and chairs central bank meetings.
- 3771 • **Deputy Governors:** Two deputy governors that primarily oversee departments (Monetary
3772 department; Financial Markets and Resolution Department), but can also serve as the acting
3773 governor if necessary.
- 3774 • **Other members:** Four other board members that manage other departments within the
3775 CNB.

3776 Meeting Structure:

- 3777 • **Frequency:** Regular meetings are held eight times annually to assess economic and mone-
3778 tary conditions.
- 3779 • **Additional Meetings:** Held during extraordinary economic developments.

3780 Manual Annotation

3781 Annotators

- 3782 • Marc-Alain Adjahi

3783 **Annotation Guide** Our annotation guide is inspired by the work of [Shah et al. \[2023\]](#). We modify
3784 their annotation guide further, keeping in mind the changing global macroeconomic dynamics. Also,
3785 we have two additional categories (Forward Looking or Not Forward Looking and Certain or
3786 Uncertain) of labels other than monetary policy stance.

3787 The annotation guide for monetary policy stance is summarized in Table 106. It is built by dividing
3788 each target sentence into six defined categories: Economic Status, Currency Stability, Inflation Rates,
3789 Interest Rates, External Shocks, and Wage Dynamics.

- 3790 • *Economic Status:* A sentence pertaining to the overall performance of the economy, particu-
3791 larly focusing on fluctuations in inflation, unemployment, and economic growth.
- 3792 • *Currency Stability:* A sentence addressing changes in the exchange rate of the Czech koruna,
3793 reflecting its appreciation or depreciation and its impact on import costs and monetary
3794 conditions.
- 3795 • *Inflation Rates:* A sentence focusing on trends in inflation, especially in relation to the CNB's
3796 target range, and indicating potential monetary policy actions in response to deviations.
- 3797 • *Interest Rates:* A sentence relating to adjustments in the CNB's interest rate policy, highlight-
3798 ing decisions to raise, lower, or maintain borrowing costs as a means to influence economic
3799 activity.

- 3800 • *External Shocks*: A sentence examining the impact of external economic factors, such as
3801 shifts in international demand or economic conditions abroad, that could influence domestic
3802 stability and policy decisions.
- 3803 • *Wage Dynamics*: A sentence describing changes in wage growth relative to productivity,
3804 assessing their effects on domestic demand, inflation pressures, and the overall balance of
3805 the economy.

3806 **Examples:**

- 3807 • “Inflation would be curbed by the strengthening of the Czech koruna vis-à-vis the euro,
3808 pulling import prices down and tightening monetary conditions in the economy.”
3809 **Dovish**: This example illustrates a scenario where easing conditions support a dovish stance.
- 3810 • “At the close of the meeting, the Board decided unanimously to lower the CNB two-week
3811 repo rate by 0.50 percentage points to 4.75%.”
3812 **Dovish**: The decision to lower rates is indicative of a dovish monetary policy approach.
- 3813 • “The risks presented by the forecast had accumulated on the inflation side, with a rise
3814 in money supply and fiscal deficit potentially sending a stronger demand impulse than
3815 anticipated.”
3816 **Hawkish**: This sentence demonstrates concerns over overheating, suggesting a hawkish
3817 response.
- 3818 • “At the close of the meeting, the Board decided unanimously to raise the CNB two-week
3819 repo rate by 0.25 percentage points to 5.25%.”
3820 **Hawkish**: The rate increase reflects a move towards tighter policy, characteristic of a
3821 hawkish stance.
- 3822 • “The actual data did not confirm a substantial weakening, and money supply growth con-
3823 tinued to reach a figure of about 12% year-on-year, supporting a stable rise in household
3824 consumption.”
3825 **Neutral**: The stable economic indicators give rise to a neutral policy outlook.
- 3826 • “The next bank board meeting on monetary and economic development will take place on
3827 28 January 1999.”
3828 **Irrelevant**: This sentence solely provides scheduling information and does not offer any
3829 insights into monetary policy.

Table 106: Czech National Bank Annotation Guide

Category	Hawkish	Dovish	Neutral	Irrelevant
Economic Status	When inflation increases, unemployment decreases, or economic growth is projected high, indicating a potential overheating of the economy.	When inflation decreases, unemployment increases, or economic growth is projected as low.	When economic indicators remain stable.	Sentence is not relevant to monetary policy.
Currency Stability	When the Czech koruna appreciates, reducing import costs, signaling a stronger economy, and potential tightening of policy.	When the Czech koruna depreciates, increasing import costs, potentially signaling a need for a looser monetary policy to stabilize the economy.	When the koruna exchange rate remains relatively stable over a period, without sharp fluctuations.	Sentence is not relevant to monetary policy.
Inflation Rates	When inflation rises above the CNB's target range, requiring tighter policy to anchor expectations.	When inflation trends downward, providing room for looser monetary policy to support growth.	When inflation rates remain stable within a target range.	Sentence is not relevant to monetary policy.
Interest Rates	When interest rates are raised, signaling the CNB's intent to curb inflation by making borrowing more expensive.	When interest rates are lowered, indicating the CNB is encouraging borrowing and investment by reducing the cost of credit.	When interest rates remain unchanged from previous levels.	Sentence is not relevant to monetary policy.
External Shocks	When external demand surges, driven by recovery in key markets, potentially requiring tighter policy to avoid overheating.	When external shocks, such as a downturn in Germany or an EU recession, weaken Czech exports, calling for monetary easing to stabilize growth.	When external economic conditions remain stable, causing no immediate impact.	Sentence is not relevant to monetary policy.
Wage Dynamics	When wage growth accelerates beyond productivity levels, risking inflation and higher costs for businesses.	When wage growth slows, signaling weakening domestic demand and easing inflationary pressures.	When wage growth aligns with productivity, posing no immediate risk to inflation or demand.	Sentence is not relevant to monetary policy.

3830 The annotation guide for whether the sentence has Forward Looking information or not is summarized in the Table 107 and whether the committee is talking about the sentence with Certainty
3831 attest to it or not in Table 108.
3832

Table 107: Czech National Bank Annotation Guide to identify whether a sentence has forward-looking information or not.

Label	Description	Example
Forward Looking	Projections of economic data and the geopolitical landscape. Comments about what is expected to come in future meetings or in future events.	“The forecast expected headline inflation to increase this year from its current very low but positive level.”
Not Forward Looking	Reflections of previous economic data and the geopolitical landscape. Comments about events that have already happened in the past.	“Average wage growth had risen to 2.2 in 2015 q1, slightly above the current forecast.”

Table 108: Czech National Bank Annotation Guide to identify whether a sentence echoes certainty or not.

Label	Description	Example
Certain	References data points or events that have already happened in the past. If the sentence references the future, it uses that reflect certainty, such as “will.”	“The labour market was meanwhile still assessed as strongly overheated.”
Uncertain	Projections of the future that do not reflect certainty. Key words include “may” and “should.”	“In light of recent events, the still unclear form of Brexit could be added to this uncertainty.”

3833 NeurIPS Paper Checklist

3834 1. Claims

3835 Question: Do the main claims made in the abstract and introduction accurately reflect the
3836 paper's contributions and scope?

3837 Answer: [Yes]

3838 Justification: Abstract (on page 1) and introduction (Section 1) accurately reflect the paper's
3839 contributions and scope.

3840 Guidelines:

- 3841 • The answer NA means that the abstract and introduction do not include the claims
3842 made in the paper.
- 3843 • The abstract and/or introduction should clearly state the claims made, including the
3844 contributions made in the paper and important assumptions and limitations. A No or
3845 NA answer to this question will not be perceived well by the reviewers.
- 3846 • The claims made should match theoretical and experimental results, and reflect how
3847 much the results can be expected to generalize to other settings.
- 3848 • It is fine to include aspirational goals as motivation as long as it is clear that these goals
3849 are not attained by the paper.

3850 2. Limitations

3851 Question: Does the paper discuss the limitations of the work performed by the authors?

3852 Answer: [Yes]

3853 Justification: In Section 8 and Section A.4.

3854 Guidelines:

- 3855 • The answer NA means that the paper has no limitation while the answer No means that
3856 the paper has limitations, but those are not discussed in the paper.
- 3857 • The authors are encouraged to create a separate "Limitations" section in their paper.
- 3858 • The paper should point out any strong assumptions and how robust the results are to
3859 violations of these assumptions (e.g., independence assumptions, noiseless settings,
3860 model well-specification, asymptotic approximations only holding locally). The authors
3861 should reflect on how these assumptions might be violated in practice and what the
3862 implications would be.
- 3863 • The authors should reflect on the scope of the claims made, e.g., if the approach was
3864 only tested on a few datasets or with a few runs. In general, empirical results often
3865 depend on implicit assumptions, which should be articulated.
- 3866 • The authors should reflect on the factors that influence the performance of the approach.
3867 For example, a facial recognition algorithm may perform poorly when image resolution
3868 is low or images are taken in low lighting. Or a speech-to-text system might not be
3869 used reliably to provide closed captions for online lectures because it fails to handle
3870 technical jargon.
- 3871 • The authors should discuss the computational efficiency of the proposed algorithms
3872 and how they scale with dataset size.
- 3873 • If applicable, the authors should discuss possible limitations of their approach to
3874 address problems of privacy and fairness.
- 3875 • While the authors might fear that complete honesty about limitations might be used by
3876 reviewers as grounds for rejection, a worse outcome might be that reviewers discover
3877 limitations that aren't acknowledged in the paper. The authors should use their best
3878 judgment and recognize that individual actions in favor of transparency play an impor-
3879 tant role in developing norms that preserve the integrity of the community. Reviewers
3880 will be specifically instructed to not penalize honesty concerning limitations.

3881 3. Theory assumptions and proofs

3882 Question: For each theoretical result, does the paper provide the full set of assumptions and
3883 a complete (and correct) proof?

3884 Answer: [NA]

Justification: The paper is submitted to the Dataset and Benchmark Track, and we have conducted extensive experiments to back any claim we make.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: We believe we do one of the most extensive experiments and reporting on experiments and the annotation process. The details on the models experimented with are provided in Section 4.1, Appendix C.2, and Appendix C.3. Prompts are provided in Appendix C.1. Instructions given to annotators are provided in Appendix B.6, while the interface used is provided in B.7. The complete list is too long to be included here, so we would like to direct you to the Table of Contents on Page 14.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
 - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
 - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
 - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
 - (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in

some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: Page 1 provides URLs to both our Code, Model, and Dataset pages. The Model and Dataset pages also provide code on how to use them. We go above and beyond to create a website for the non-technical audience to use our work.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. Experimental setting/details

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: The details on dataset splits and other settings are provided in Section 4.1, Appendix C.2, and Appendix C.3. Prompts are provided in Appendix C.1. Hyperparameter details are provided in Appendix C.8.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. Experiment statistical significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: Every experiment is conducted over multiple seeds, and we report the standard deviation of metrics to provide understanding of the statistical significance of the experiments. Additionally, whenever we run regression or correlation, we provide the p-value.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. Experiments compute resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: We provide extensive details on packages used, hardware used, and outside services used in Appendix C.3.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. Code of ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: We provide an author statement in the Appendix A.1, Disclaimer in Appendix A.3, Ethical Consideration in A.4, and Scraping Policy Consideration in A.6.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. Broader impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: In Section 8.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [Yes]

Justification: We provide the dataset clause in Appendix A.5. For models, we fine-tune existing models and cite them.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: We cite all the models used in Section 4.1. We collect the original new data with Scraping Policy Consideration in Appendix A.6.

Guidelines:

- The answer NA means that the paper does not use existing assets.

- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, paperswithcode.com/datasets has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. New assets

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [Yes]

Justification: Appendix A.2 provide details on this extensively.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. Crowdsourcing and research with human subjects

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [Yes]

Justification: Instructions given to annotators are provided in Appendix B.6, while the interface used is provided in B.7.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. Institutional review board (IRB) approvals or equivalent for research with human subjects

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: We do not perform any study that requires it.

4147
4148
4149
4150
4151
4152
4153
4154
4155
4156
4157
4158
4159
4160
4161
4162
4163
4164
4165
4166
4167
4168
4169

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: We only use LLM for performance benchmarking.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>) for what should or should not be described.