forecasting

MDPI

Article

When Does Central Bank Communication Matter? Textual Information, Dynamics, and Regularization

Igor Barbosa de Andrade Duarte and Márcio Poletti Laurini *

Faculty of Economics, Administration and Accounting of Ribeirão Preto, University of São Paulo, Ribeirão Preto 14040-905, SP, Brazil

* Correspondence: laurini@fearp.usp.br

Highlights

What are the main findings?

- Textual information from central bank communication improves forecast accuracy for interest rate changes when combined with policy dynamics and regularization techniques.
- Forecast gains are concentrated around policy regime shifts, where persistence-based benchmarks generate large forecast errors.

What are the implications of the main findings?

- Regularization and signal extraction are essential to transform high-dimensional textual data into economically relevant predictive signals.
- Central bank communication contains forward-looking information that is most valuable for forecasting, asset pricing, and risk management during periods of policy adjustment.

Abstract

This paper examines when textual information from central bank communication improves forecasts of policy rate changes. Using the minutes of the Brazilian Central Bank's Monetary Policy Committee (Copom), we study whether textual content helps predict changes in the Selic target rate between consecutive meetings. The minutes are encoded using dense sentence-level embeddings, and low-dimensional textual factors are extracted via principal component analysis estimated exclusively on the training sample to prevent look-ahead bias. Predictive performance is assessed out of sample using an expanding-window backtesting framework and compared against standard forecasting benchmarks, including persistence and random-walk specifications, linear autoregressive models, regularized regressions, and state-space models estimated via the Kalman filter. We find that text-based predictors perform poorly when used in isolation but deliver meaningful forecast improvements when combined with short-run dynamics and regularization. These gains are economically relevant and arise primarily in episodes associated with policy rate adjustments, whereas simple persistence-based forecasts remain difficult to outperform during rate-hold periods. Overall, the results indicate that central bank communication contains forward-looking information that is valuable for forecasting policy changes, but that this information is sparse, episodic, and best extracted through disciplined regularization and dynamic modeling rather than purely cross-sectional textual signals.

Keywords: text-as-data; machine learning in macroeconomics; sentence embeddings; Kalman filtering; forecast evaluation

Academic Editor: Sonia Leva

Received: 5 February 2026
Revised: 12 March 2026
Accepted: 13 March 2026
Published: 16 March 2026

1. Introduction

Central bank communication is a core component of the monetary policy transmission mechanism. Beyond the policy decision itself, central banks convey assessments of the macroeconomic outlook, the balance of risks, and their likely reaction function to future shocks [1]. These signals shape expectations across financial markets, influencing the pricing of fixed-income securities, currencies, and other interest-sensitive assets [2]. As a result, understanding whether—and when—central bank communication contains forward-looking information about future policy actions is of first-order importance for both macroeconomic analysis and asset pricing.

In Brazil, the minutes of the Monetary Policy Committee (Copom) constitute the main vehicle through which the Central Bank communicates its interpretation of economic conditions and its policy stance [2]. While the policy rate decision is announced immediately after each meeting, the minutes provide a richer and more nuanced narrative of the underlying deliberations. Market participants routinely scrutinize these documents to infer whether future adjustments to the Selic target rate are more or less likely. Yet, despite their prominence in practice, systematic evidence on the strictly out-of-sample predictive content of Copom minutes for subsequent policy decisions remains limited.

This paper addresses a simple but economically meaningful question: given the recent history of policy decisions and the textual content of the minutes released after meeting t, can one forecast changes in the Selic target rate at meeting $t+1$ more accurately than with simple time-series benchmarks? Importantly, we frame this question in a purely predictive and real-time setting, abstracting from in-sample explanatory power or ex post narrative coherence. Our focus is on whether information embedded in central bank communication improves forecasts relative to strong benchmarks that reflect the well-known inertia of monetary policy.

We define the forecasting target as $y_{t+1} = 100\,(\text{Selic}_{t+1} - \text{Selic}_t)$, measured in basis points, and evaluate one-step-ahead forecasts using an expanding-window backtesting protocol that mimics recursive forecasting with re-estimation as new information becomes available. This design closely mirrors the decision environment faced by investors, analysts, and policymakers and rules out look-ahead bias by construction.

The empirical setting presents nontrivial challenges. Changes in the Selic rate are discrete and characterized by a substantial mass at zero, reflecting frequent rate-hold decisions interspersed with relatively rare adjustments, typically in increments of 25 or 50 basis points. This feature makes naïve benchmarks, such as persistence or random walks, remarkably competitive, particularly during periods of policy inertia. At the same time, standard loss functions such as the root mean squared error (RMSE) can be dominated by a small number of large forecast errors associated with rate changes. This motivates an evaluation strategy that combines global accuracy measures with regime-specific diagnostics that distinguish between maintenance and adjustment episodes.

From a forecasting perspective, incorporating textual information into predictive models raises two intertwined challenges: extracting a weak, noisy signal from high-dimensional representations, and combining that signal with strong but mechanical time-series dynamics. Both challenges are particularly acute in monetary policy settings, where the target variable is discrete, heavily concentrated at zero, and dominated by persistence.

Textual representations of central bank communication, such as embeddings derived from policy minutes, are inherently high dimensional and statistically ill-posed relative to the available sample size. Even after dimension reduction, the resulting factors are estimated with error and may be highly correlated. In short samples, naïvely treating these factors as standard regressors risks overfitting and unstable forecasts, especially out of sample. Regularization is therefore not a modeling choice of convenience but a necessary

mechanism for disciplined signal extraction. By shrinking coefficients toward zero when the data provide weak evidence, penalized estimators explicitly trade variance for bias, stabilizing forecasts and preventing spurious amplification of textual noise.

At the same time, purely text-based forecasts ignore a critical benchmark: past policy actions already embed a strong, low-variance signal about future decisions. In financial terms, persistence-based forecasts act as a dominant factor with high explanatory power during most periods, particularly when policy is on hold. The relevant forecasting problem is therefore not whether text predicts policy decisions in isolation but whether it adds incremental information conditional on the prevailing policy stance. This naturally leads to a forecast-combination interpretation: dynamic terms capture the baseline expectation implied by recent decisions, while textual factors provide an adjustment that reflects forward-looking communication about risks, balance of considerations, and potential regime shifts.

Regularization plays a central role in governing this combination. In penalized dynamic models, the persistence component is allowed to dominate when textual evidence is weak or inconsistent, effectively reverting forecasts toward a conservative benchmark. Conversely, when textual signals are persistent and aligned with subsequent outcomes, regularization allows them to enter gradually and systematically, rather than through large, unstable coefficient estimates. In this sense, regularization functions as an implicit weighting scheme, endogenously determining when and how much weight should be placed on communication relative to past actions.

Viewed through this lens, models that combine dynamics and text under regularization are best understood as signal-extraction devices rather than structural representations of policy behavior. They filter noisy textual information through the lens of historical policy inertia, producing forecasts that are robust in tranquil periods yet responsive during transitions. This is particularly valuable in financial applications, where forecast errors are most costly precisely when policy regimes shift and asset prices reprice rapidly.

More broadly, this framework aligns central bank communication with standard forecasting principles: high-dimensional predictors should be shrunk, weak signals should be filtered rather than amplified, and new information should complement—not override—strong baseline predictors. Regularization, signal extraction, and forecast combination are therefore not ancillary features of the empirical design but the core mechanisms through which textual information becomes economically relevant for forecasting monetary policy outcomes.

To transform the textual content of the minutes into quantitative predictors, we represent each document using dense sentence-level embeddings. Given the high dimensionality of these representations relative to the limited number of policy meetings, we extract low-dimensional textual factors via principal component analysis (PCA), estimated strictly on the training sample at each iteration of the backtest. This approach yields parsimonious regressors that summarize semantic variation, such as shifts in tone, emphasis, and thematic focus—while mitigating overfitting and preserving the integrity of the out-of-sample evaluation.

To operationalize textual data for predictive analysis, this study leverages the paraphrase-multilingual-MiniLM-L12-v2 model as its core semantic encoder. This model is a distilled, transformer-based architecture specifically engineered to generate high-quality, language-agnostic sentence embeddings. Unlike generative large language models (LLMs), its design is optimized for representation learning: it converts input text into a dense 384-dimensional vector space where semantic similarity is preserved as geometric proximity. We selected this model for its trifecta of methodological advantages essential to forecasting: its multilingual proficiency, enabled by knowledge distillation across over

50 languages, ensures robust performance on diverse global sources without intermediary translation; its compact efficiency (12-layer encoder) allows for the scalable processing of longitudinal text corpora; and its dedicated fine-tuning on paraphrase identification explicitly optimizes it for capturing complex semantic equivalences and thematic shifts. By transforming unstructured discourse into a structured time series of semantic embeddings, this model provides the foundational linguistic signal for our subsequent forecasting pipeline, bridging the gap between qualitative text and quantitative prediction.

Model selection follows a transparent and incremental strategy. We begin with two widely used benchmarks in settings with discrete outcomes and strong persistence: a random walk in differences, which formalizes rate maintenance as the baseline forecast, and a persistence model that extrapolates the most recent policy change. We then consider a text-only specification to assess whether communication can substitute policy dynamics. Our main specification combines both sources of information by augmenting a lagged dependent variable with textual factors. Given the short sample and correlated regressors, this model is estimated using Ridge regularization, with Elastic Net considered as a robustness check. Finally, we explore a state-space SARIMAX formulation estimated via maximum likelihood through the Kalman filter, connecting the exercise to the econometric literature on dynamic models and offering an alternative to penalized regressions.

We place particular emphasis on controlling overfitting and temporal leakage. Dimensionality reduction, hyperparameter selection, and model estimation are all restricted to the training sample, and forecast comparisons are conducted using formal predictive accuracy tests. The resulting workflow—from PDF text extraction to recursive estimation and automated generation of tables and figures—is fully reproducible and designed to facilitate auditability.

The empirical results reveal a pronounced asymmetry in forecast performance that helps clarify when—and how—central bank communication matters for prediction. Simple persistence-based benchmarks remain difficult to outperform in directional metrics, reflecting the dominance of rate-hold decisions and the strong inertia of the policy cycle. Models that rely exclusively on textual information fail to deliver out-of-sample gains, indicating that communication does not substitute for the information embedded in recent policy actions. Instead, textual signals appear to be incremental and conditional on the underlying dynamics of the policy process.

Forecast improvements emerge only when textual factors are integrated with short-run policy dynamics and disciplined regularization. Specifications that combine lagged policy changes with text-derived factors and employ shrinkage achieve lower forecast errors in terms of RMSE, with gains concentrated in meetings associated with policy rate adjustments. Regime-specific diagnostics reinforce this pattern: while persistence performs well during maintenance episodes, text-augmented and regularized models tend to reduce the magnitude of forecast errors precisely during transitions of the monetary policy cycle.

This evidence highlights the dual role of regularization in extracting predictive content from central bank communication. First, shrinkage stabilizes estimation in short samples with noisy and potentially collinear textual factors. Second, it effectively filters text-based information, allowing forward-looking signals to influence forecasts only when they meaningfully complement policy dynamics. From a practical forecasting perspective, these transition episodes occur precisely when mispricing risks are largest and the marginal value of improved predictions is highest for investors and risk managers.

More broadly, the results suggest that the predictive value of central bank communication is not uniform over time but state-dependent and episodic. By jointly emphasizing dynamics and regularization, this paper contributes to the text-as-data literature by showing that communication matters not as a standalone predictor but as a selectively informa-

tive signal that becomes relevant during periods of policy adjustment—when forecasting accuracy is most valuable.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the data, text processing pipeline, models, and backtesting protocol. Section 4 presents the empirical results, including regime-specific diagnostics and forecast comparison tests. Section 5 concludes by summarizing the findings, discussing limitations, and outlining avenues for future research.

2. Literature Review

Under inflation-targeting regimes, monetary policy is transmitted primarily through expectations rather than solely through contemporaneous policy actions. Seminal theoretical contributions emphasize that the ability of central banks to shape beliefs about the future path of policy rates is central to policy effectiveness in environments with nominal rigidities and forward-looking agents [3]. Within this framework, central bank communication is not merely an explanatory device but a substantive policy instrument that influences private-sector expectations, financial conditions, and asset prices.

A growing empirical literature documents that official communications—such as policy statements, minutes, speeches, and inflation reports—contain information that markets actively process when forming expectations about future monetary policy [1]. Early evidence shows that monetary policy announcements affect long-term interest rates primarily through changes in expected future short rates rather than through contemporaneous policy surprises [4]. Subsequent work demonstrates that forward guidance, whether explicit or implicit, moves yield curves, exchange rates, and risk premia by revealing information about the central bank's reaction function and economic outlook [5–7].

Recent advances in text-as-data methods have enabled a more systematic quantification of central bank communication [8]. Studies using textual analysis of policy statements and minutes find that changes in tone, topics, and emphasis are informative about future policy actions and macroeconomic conditions [9,10] and also affect key financial variables [11]. These approaches highlight that communication can affect markets independently of the policy rate decision itself, reinforcing the view that communication constitutes a distinct and powerful policy channel.

In emerging market economies, where credibility, inflation persistence, and exposure to external shocks are particularly salient, communication may play an even more prominent role. In the Brazilian case, the Central Bank relies heavily on post-meeting statements, detailed minutes of the Monetary Policy Committee (Copom), and periodic inflation reports to convey qualitative assessments of domestic and global conditions, risk balances, and policy trade-offs [12,13]. Market participants routinely scrutinize these documents to infer whether the authority is leaning toward tightening, easing, or maintaining the policy stance in subsequent meetings.

Despite the recognized importance of communication, relatively little is known about it's strictly out-of-sample predictive content for future policy decisions, particularly when evaluated against strong time-series benchmarks that reflect the inherent inertia of monetary policy. Much of the existing evidence focuses on in-sample correlations, event-study reactions, or explanatory regressions, which may overstate the informational value of communication in real-time forecasting environments. Moreover, the discrete and infrequent nature of policy rate changes poses additional challenges for assessing predictive gains in a statistically and economically meaningful way.

Evidence for the Brazilian case indicates that the tone of central bank communication affects the yield curve and asset prices and suggests that the policy statement anticipates part of the information later conveyed in the minutes, thereby reducing their strictly new

informational content, as documented by [14]. Central bank communication has been shown to contain predictive information for movements in the Brazilian term structure of interest rates. In particular,refs. [13,15] document that communication-related signals help explain futures rate dynamics, while ref. [12] provides evidence of economically meaningful out-of-sample predictive power.

The literature also emphasizes the distinction between surprises associated with the policy decision itself, the target component, and revisions to the expected future path of policy rates—the path component. The latter is more directly linked to the role of communication in shaping expectations about future monetary policy, as discussed by [16].

To examine this informational role empirically, textual communication must be transformed into quantitative variables. Traditional text-as-data approaches rely on bag-of-words representations, dictionaries, and word counts, which offer transparency and low computational cost but are limited when meaning depends on context, negation, intensity, and semantic structure, as emphasized by [17]. More recently, embeddings and neural language models have enabled the representation of sentences and documents as dense vectors that capture semantic similarity and latent factors. Emerging evidence suggests that embeddings tailored to the domain of central bank communication—particularly when combined with fine-tuning strategies and dimensionality reduction to address high dimensionality—can outperform purely lexical measures in certain economic applications, as illustrated in [18–20].

Empirical evidence suggests that textual information can predict, or, more cautiously, help organize information about—a range of economic outcomes, including policy decisions and expectations, financial market reactions, and, in some settings, macroeconomic variables. For the United States, ref. [9] show that measures extracted from FOMC minutes are systematically related to movements in interest rate futures, exchange rates, and equity prices, indicating that more detailed documents may convey additional signals about the economic outlook and the balance of risks. In out-of-sample evaluations, however, predictive gains tend to be heterogeneous across types of communication and institutional designs and depend critically on measurement choices and informational constraints. These findings motivate the present study's focus on extracting low-dimensional textual factors from Copom minutes—using embeddings and principal component analysis—and on assessing, within a rigorous out-of-sample forecasting framework, whether such factors contain predictive information for subsequent changes in the policy interest rate.

In Brazil, the institutional design of Copom communication provides a particularly useful setting for discussing incremental information. Policy statements are released immediately after each meeting, while the minutes are published a few days later. Although other news may arrive in the interim, this sequencing reduces temporal overlaps and helps structure the flow of information available to markets. The literature exploits this proximity in timing to distinguish information that is priced immediately after the decision and statement from what is genuinely new in the minutes. In this context, timing restrictions in the construction of textual measures become central: textual scoring should rely exclusively on information available at the time of release in order to minimize contamination from subsequent data revisions or later disclosures, as proposed by [2] and further discussed by [21].

This literature also emphasizes that the relative informational role of policy statements and minutes evolves over time. Changes in textual format and content, particularly after 2016—when policy statements began to incorporate more explicit discussion of the economic outlook and risk balance—affect intertemporal comparability and the extent to which minutes contain genuinely new information. These changes also highlight inherent limitations of official documents, including repetition, standardization, and the selective

disclosure of deliberative content. For the purposes of this study, these considerations motivate restricting the sample to the contemporary communication regime and treating the statement–minutes window as an operational boundary for what can plausibly be interpreted as new information relevant for forecasting, in line with the evidence and discussions in [21,22].

From a methodological perspective, the macroeconomic forecasting literature underscores that textual information adds value only when evaluated in a genuinely out-of-sample setting and under careful controls against temporal leakage, a concern that is amplified by short samples and high-dimensional predictors. In such environments, regularization (e.g., the ridge, lasso, elastic net methods), dimensionality reduction, and time-series-appropriate validation schemes are essential tools for mitigating overfitting and parameter instability. The foundations of regularization date back to [23], with subsequent developments in lasso and elastic net by [24,25]. Best practices and trade-offs in text-as-data applications are systematically discussed by [17]. Ultimately, formal comparisons of predictive accuracy remain the relevant criterion for assessing whether text delivers meaningful gains over parsimonious benchmarks, as emphasized by [26].

For central bank communication in particular, forecasting exercises are especially demanding because text may simultaneously convey signal and noise, and its predictive contribution can depend sensitively on document type and institutional design. Recent evidence suggests that incorporating multiple textual sources can, in some cases, degrade out-of-sample performance by introducing additional noise through excessive granularity. For Brazil, existing applications indicate that the gains from machine learning methods using central bank communications are heterogeneous [12] and depend on well-defined forecast competitions and evaluation windows. Moreover, studies that construct modern textual measures—including tone, novelty, and domain-specific fine-tuning—consistently stress that empirical validity hinges on strict alignment with information timing and disclosure constraints, as discussed in [27,28].

Against this backdrop, this paper contributes by implementing a fully reproducible pipeline that transforms Copom minutes into dense embeddings, extracts low-dimensional textual factors via PCA, and incorporates them into dynamic and regularized forecasting models. These models are evaluated through recursive backtests and formal forecast accuracy tests, allowing us to characterize when—and under which informational and institutional constraints—textual information adds incremental predictive value for future policy rate changes within the Brazilian communication framework.

Regularization has long been recognized as essential for improving out-of-sample performance in predictive settings characterized by limited sample sizes, persistent regressors, and collinearity. Early contributions emphasize the bias–variance trade-off inherent in forecasting problems, motivating shrinkage estimators that reduce estimation error at the cost of controlled bias [23,24]. In time-series and financial forecasting, regularization plays a dual role: stabilizing parameter estimates and implicitly imposing structure on the dynamics of the predictive relationship [29,30].

In macro-finance applications, shrinkage methods such as Ridge and Elastic Net have been shown to outperform unregularized benchmarks when predictors are numerous and weakly informative [31,32]. From a state-space perspective, regularization can be interpreted as a prior on parameter evolution or signal smoothness, linking penalized regression to filtering and Bayesian updating [33,34]. This connection provides a unifying framework in which static shrinkage and dynamic filtering are viewed as alternative mechanisms for controlling overfitting while extracting persistent predictive signals.

The use of textual data in economics and finance has expanded rapidly with advances in natural language processing. A central challenge in this literature is dimensionality

reduction: raw text representations are extremely high-dimensional and noisy relative to the available time-series length. As a result, most successful applications rely on the construction of low-dimensional summaries that act as sufficient statistics for forecasting tasks [8,35].

Early approaches relied on dictionary-based methods and topic models [36,37], while more recent work uses dense embeddings derived from neural language models [38]. Regardless of the representation, predictive performance typically depends on extracting a small number of latent factors that concentrate the relevant variation in the text. Factor-based approaches using principal components or related techniques are common in macro-finance forecasting, where they provide a disciplined way to map high-dimensional predictors into a stable forecasting space [29].

3. Data and Methodology

3.1. Data, Sample, and Target Variable Definition

The unit of observation in this study is the meeting of the Monetary Policy Committee (Copom). Let Selic_t denote the target Selic rate set at meeting t, observed immediately following the policy decision. The objective is to forecast, using information associated with meeting t, including the textual content of the minutes released subsequently, the change in the Selic rate at the next meeting.

The target variable is defined as the change in the policy rate between consecutive decisions, measured in basis points:

$$y_{t+1} \equiv \Delta\text{Selic}_{t+1} = 100 \cdot (\text{Selic}_{t+1} - \text{Selic}_t), \tag{1}$$

so that y_{t+1} takes discrete values (e.g., 0, $\pm$25, $\pm$50 bps).

The sample construction follows an operational rule whereby only minutes from the 200th Copom meeting onward are included, and thus the sample starts with the meeting on 19–20 July 2016. From this point, the Central Bank adopts a standardized textual structure consistent with the contemporary format of the minutes. This restriction mitigates heterogeneity in writing style and document organization and enhances the comparability of textual information extraction over time—a consideration also emphasized by [21]. Under this criterion, the final dataset comprises $T = 75$ meetings. The last meeting included was the 274th, held on 4 and 5 November 2025.

The predictive evaluation employs a minimum training window of size $t_0 = 50$, yielding $N = T - t_0 = 25$ out-of-sample forecasts in the test period. The forecasting protocol follows an expanding-window scheme and is described in detail in Section 3.13.

3.2. Text Extraction, Preprocessing, and Chunking

For each meeting t, the PDF file containing the Copom minutes is collected, and its content is converted into raw text. Monetary policy documents are typically long and detailed, which raises two practical constraints: (i) computational limits associated with large-scale text processing, and (ii) context-length restrictions imposed by modern embedding models. To address these constraints while still exploiting the full informational content of each document, we adopt a chunkingprocedure following the general approach discussed in [39].

After an initial preprocessing stage—consisting of standard steps such as removal of formatting artifacts, normalization of whitespace, and elimination of non-textual elements—the text of each set of minutes is partitioned into C_t contiguous segments (chunks) of approximately constant length, measured either in characters or tokens. Let $\mathcal{T}_{t,c}$ denote the text associated with chunk c from meeting t.

This chunking step serves three complementary purposes. First, it renders the embedding extraction computationally feasible by ensuring that each textual unit fits within the model's context window. Second, it mitigates the risk of information loss due to truncation of long documents. Third, it allows embeddings to capture local semantic context, which is particularly relevant in policy documents where different sections may convey distinct signals regarding economic conditions, risks, and policy inclinations.

3.3. Mathematical Formalization of Textual Embeddings

Let $\mathcal{V}$ denote a finite vocabulary and let $\mathcal{T}$ denote the space of finite-length texts over $\mathcal{V}$, where a text may correspond to a sentence, paragraph, or document. A textual embedding is defined as a mapping

$$f : \mathcal{T} \to \mathbb{R}^d,$$

which associates each text $T \in \mathcal{T}$ with a vector $f(T)$ in a d-dimensional continuous vector space.

The embedding function $f(\cdot)$ is typically parameterized by a function class $\mathcal{F} = \{f_\theta : \theta \in \Theta\}$ and obtained as the solution to an optimization problem defined on a large external corpus. The training objective is designed so that the geometry of the embedding space reflects semantic and syntactic regularities of natural language. In particular, for any two texts $T, T' \in \mathcal{T}$, the distance or similarity between $f(T)$ and $f(T')$ is intended to approximate a latent semantic proximity measure,

$$\|f(T) - f(T')\| \approx \delta(T, T'),$$

where $\delta(\cdot, \cdot)$ is an unobserved semantic distance induced by language usage.

In neural embedding models, the function f_θ is implemented as a composition of tokenization, contextual encoding, and pooling operators. Let $T = (w_1, \ldots, w_n)$ denote a sequence of tokens. The embedding is constructed as

$$f_\theta(T) = \mathcal{P}(g_\theta(w_1, \ldots, w_n)),$$

where g_θ is a contextual encoder that maps token sequences into a sequence of hidden states in $\mathbb{R}^d$, and $\mathcal{P}$ is a pooling operator that aggregates token-level representations into a single fixed-dimensional vector. Common pooling choices include averaging, weighted averaging, or selection of a designated summary token.

From a statistical perspective, the embedding $f(T)$ can be interpreted as a nonlinear transformation of the original discrete text into a continuous feature vector of fixed dimension d, which does not grow with text length and therefore facilitates downstream statistical analysis. Individual coordinates of $f(T)$ generally lack direct interpretability; instead, meaning is encoded in the relative geometry of the embedding space. In this representation space, texts can be compared and combined using standard linear-algebraic operations, and the embedding serves as a low-dimensional proxy for the informational content of the text. Its adequacy is evaluated through performance in downstream tasks such as prediction, classification, or inference rather than by inspection of individual components. Applications of embedding-based representations in economic and financial contexts are discussed in [17].

Because the COPOM minutes are too long to be processed directly by standard embedding models, we apply a chunking procedure in which each document is partitioned into segments $\mathcal{T}_{t,c}$. Each chunk is then mapped into a dense numerical vector through the embedding function.

To obtain a single representation for each Copom meeting, the information contained in its chunks is aggregated via a simple average:

$$\bar{e}_t = \frac{1}{C_t} \sum_{c=1}^{C_t} e_{t,c}, \tag{2}$$

where $\bar{e}_t \in \mathbb{R}^d$ denotes the document-level embedding associated with the minutes of meeting t.

From a conceptual standpoint, the embedding vector e_t can be interpreted as a low-dimensional summary statistic that aggregates the semantic content of the document. It captures information about the overall tone, the topics discussed, relative emphasis across themes, and patterns of lexical co-occurrence without imposing an explicit economic or institutional structure. In this sense, embeddings can be viewed as an empirical approximation to an (approximately) sufficient statistic for forecasting, consistent with the literature that emphasizes optimal information compression for prediction rather than direct causal interpretability of regressors.

This interpretation connects to the literature on sufficient forecasting [35] and forecast combinations [40], which shows that high-dimensional information sets can often be summarized by a small number of latent factors without loss of predictive content. Within this framework, embeddings act as data-driven factors extracted from text, analogous to latent predictors in dynamic factor models or sufficient dimension reduction methods, but learned from unstructured language rather than numerical panels.

Compared to dictionary-based approaches [8], which map text into pre-specified semantic categories through word counts or tone scores, embeddings trade transparency for flexibility. Dictionary methods are interpretable but may fail to capture contextual meaning, negation, intensity, and interactions across words. Topic models such as Latent Dirichlet Allocation (LDA) [36] or Structural Topic Models (STM) [41] provide an intermediate approach, producing interpretable topics but relying on strong assumptions about topic structure and specification choices.

Embeddings instead represent semantic regularities in a continuous vector space learned from large external corpora. Although they do not explicitly model document hierarchy or deliberative structure, they can be interpreted as a flexible aggregation of multiple latent semantic signals. In forecasting applications, this representation resembles an implicit forecast combination: the vector e_t pools information from many overlapping semantic dimensions, each of which may carry weak predictive power individually but may jointly improve predictive performance.

Because the primary objective of this study is forecasting under realistic information constraints, the relevance of embeddings is evaluated through their incremental predictive content. Whether e_t provides a useful sufficient statistic is therefore an empirical question, assessed by its ability to improve out-of-sample forecasts of future policy decisions when incorporated into regularized and dynamically evaluated predictive models.

3.4. Transformer-Based Semantic Encoding for Forecasting

The present work employs the `paraphrase-multilingual-MiniLM-L12-v2` model as a dedicated semantic encoder to transform unstructured textual data into numerical embeddings suitable for temporal forecasting analysis. This model is architected upon the Transformer encoder framework [42], which utilizes a multi-head self-attention mechanism to generate contextually informed representations of input sequences. The encoder comprises a stack of twelve layers, each containing a self-attention sub-layer and a position-wise feed-forward network, stabilized by residual connections and layer normalization. This design enables the model to process an entire sentence bidirectionally, capturing long-

range linguistic dependencies and disambiguating word meanings based on full sentential context—a capability critical for accurate semantic interpretation.

The model's operational objective is not text generation but representation learning. It produces a fixed-dimensional vector ($d \in \mathbb{R}^{384}$) for an input sentence, typically via mean pooling of the final layer's token outputs. This vector functions as a dense semantic fingerprint. The model's specific pre-training via masked language modeling, followed by knowledge distillation [43] and fine-tuning on multilingual paraphrase identification tasks, explicitly optimizes its embedding space such that sentences with equivalent meanings are mapped to proximate vectors regardless of surface form or language.

For forecasting applications, this architecture presents distinct advantages. First, its multilingual capacity, developed through training on over 50 languages, ensures robust cross-lingual performance without task-specific adaptation, allowing for the consistent analysis of global text corpora. Second, its specialized training for paraphrase recognition yields embeddings with high semantic discriminability, enabling the detection of nuanced thematic shifts in discourse over time. Finally, the model's compact size (12 layers, 384-dimensional output) offers an efficient balance between representational power and computational feasibility, a practical necessity for processing large-scale, high-frequency text streams in longitudinal studies. By serving as a semantic transducer, the model converts volatile textual data into a stable, analyzable time series of embeddings, thereby providing a quantifiable linguistic signal that can be integrated with conventional econometric or machine learning forecasters to capture dynamics otherwise latent in unstructured data.

3.5. Textual Factors via PCA

The document-level embeddings e_t are high dimensional, while the number of policy meetings is relatively small. Directly including the full embedding vector as a set of regressors would therefore lead to severe statistical instability, multicollinearity, and overfitting, especially in a forecasting environment with limited time-series variation. To address these issues, we reduce dimensionality using Principal Component Analysis (PCA).

Let E denote the $T \times d$ matrix whose t-th row is $e_t^\top$. Let $P_k(\cdot)$ denote the projection operator onto the subspace spanned by the first k principal components of E. We define the vector of textual factors as

$$f_t = P_k(e_t) \in \mathbb{R}^k, \qquad k \in \{1, 2, 3, 4, 5\}. \tag{3}$$

From an econometric perspective, f_t can be interpreted as a small set of latent factors that summarize the dominant modes of variation in high-dimensional textual information.

The use of PCA is well suited to forecasting settings with short samples and highly correlated predictors. First, PCA provides an unsupervised and parsimonious representation that concentrates explanatory variation into a small number of orthogonal components. Second, by construction, PCA mitigates multicollinearity among regressors, which is particularly relevant when embeddings encode overlapping semantic information. Third, restricting attention to the leading components imposes a form of regularization that improves the bias–variance trade-off in out-of-sample prediction.

To ensure genuine out-of-sample evaluation, the computation of $P_k(\cdot)$ respects the information set available at each point in time. In each iteration of the expanding-window backtest, PCA is estimated using only the training subsample and subsequently applied to the test observation. This procedure prevents temporal leakage by ensuring that the rotation, scaling, and factor loadings implicit in PCA do not exploit future information.

A practical issue concerns the choice of the number of components k. Given the short time dimension ($T = 75$) and the objective of maintaining a parsimonious forecasting environment, we restrict attention to a small pre-specified range, $k \in \{1, 2, 3, 4, 5\}$.

This cap keeps the textual block low dimensional relative to the minimum training size in the backtest, limiting estimation variance and avoiding excessive researcher degrees of freedom.

Conceptually, k governs a bias–variance trade-off: larger values incorporate more semantic variation from the minutes but increase sensitivity to sampling noise in the estimated factor space, while smaller values impose stronger compression that may discard predictive content. Rather than relying solely on explained-variance criteria—which need not align with predictive relevance—we evaluate forecasting performance across this grid and report the sensitivity of out-of-sample accuracy to k in Section 4.

3.6. Predictive Models

To assess the informational content of textual features, we compare a sequence of predictive models ordered from the most parsimonious to the most comprehensive. In all cases, the objective is to construct one-step-ahead forecasts $\hat{y}_{t+1|t}$. We begin with simple benchmark models based solely on the history of policy decisions, followed by a baseline specification that incorporates textual information through PCA-based factors estimated by ordinary least squares. We then consider augmented specifications in which textual factors are combined with lagged policy decisions and estimated under Ridge and Elastic Net regularization. Finally, we present a dynamic specification formulated in state-space form and estimated via the Kalman filter.

3.7. Benchmarks

We adopt two simple and widely used benchmarks that are particularly relevant in forecasting environments characterized by discrete outcomes and a high frequency of policy-rate holds:

$$\text{Random walk (0 bps):} \quad \hat{y}_{t+1|t} = 0, \tag{4}$$

$$\text{Persistence:} \quad \hat{y}_{t+1|t} = y_t. \tag{5}$$

The random walk benchmark with zero change represents the hypothesis that, absent compelling new information, the most likely policy decision is to leave the policy rate unchanged. This benchmark is especially strong in environments where maintenance decisions dominate the sample and where the loss function penalizes large forecast errors symmetrically. As such, it provides a conservative baseline that any informational signal must outperform to demonstrate genuine predictive content.

The persistence benchmark captures short-run inertia in the policy decision process. In practice, monetary policy often unfolds in sequences of consecutive rate hikes or cuts, reflecting gradual adjustment, or in extended periods of stability when the policy stance is judged to be appropriate. By extrapolating the most recent change, the persistence model encodes this cyclical behavior in a minimal way. Its strong empirical performance in many settings makes it a demanding benchmark: improvements relative to persistence indicate that a model is capturing information beyond mechanical continuation of recent policy actions.

3.8. Text-Only Model (PCA(k) + OLS)

As a baseline for assessing the informational content of text in isolation, we consider a specification in which the forecast depends exclusively on textual factors extracted from the minutes associated with meeting t:

$$y_{t+1} = \alpha + \beta^\top f_t + \varepsilon_{t+1}, \tag{6}$$

where $f_t \in \mathbb{R}^k$ denotes a vector of textual factors given by the first k principal components obtained from PCA applied to text embeddings. Importantly, these factors are constructed using only information available in the training sample at each step of the expanding window. The coefficient vector $\beta \in \mathbb{R}^k$ measures the marginal contribution of each textual factor to the prediction of y_{t+1}.

This specification serves as a purely textual baseline that directly addresses the central empirical question of the paper: whether the semantic content of the minutes alone contains predictive information about subsequent policy decisions. By excluding lagged policy changes y_t, the model rules out mechanical persistence as a source of predictive gains so that any explanatory power must originate from the information embedded in the text. The text-only model also provides a benchmark against which richer specifications can be evaluated, allowing us to distinguish improvements arising from policy dynamics (via lagged outcomes) from those due to better statistical treatment of the high-dimensional textual block through regularization and dimension reduction.

Empirically, this is a demanding specification. Given the high concentration of observations at $y_{t+1} = 0$, a model relying solely on textual information faces difficulty outperforming persistence-based benchmarks in short samples. From a statistical perspective, the vector of textual factors f_t can be interpreted as a low-dimensional transformation of the original text that may approximate a sufficient statistic for forecasting y_{t+1}. Excluding lagged outcomes therefore sharpens the test of informational content: any predictive ability must arise from the extent to which the extracted textual factors summarize the relevant information set available at time t.

3.9. Main Model (PCA(k) + Ridge + lag(y_t))

The main specification combines two mechanisms that are central in the context of Copom decisions. First, policy inertia and short-run persistence: successive decisions tend to be correlated, and the distribution of policy-rate changes exhibits a large mass at zero. Second, forward-looking information embedded in communication: the minutes may signal changes in the policy stance before they materialize in the actual decision. The resulting specification is

$$y_{t+1} = \alpha + \rho\, y_t + \beta^\top f_t + \varepsilon_{t+1}, \tag{7}$$

where y_{t+1} denotes the change in the Selic target (in basis points) at meeting $t+1$, y_t is the change at the previous meeting, and $f_t \in \mathbb{R}^k$ is a vector of textual factors obtained via PCA from embeddings and available as of meeting t.

Given the high concentration of observations at $y_{t+1} = 0$, the autoregressive term y_t captures a substantial share of maintenance behavior, thereby reducing the burden placed on the textual block to explain low-frequency noise. In this formulation, textual information acts primarily as a complementary signal that becomes relevant in episodes of policy change, when forward guidance and shifts in tone may precede adjustments in the policy rate.

Even for small values of k (e.g., $k = 5$ principal components), the textual factors tend to be correlated and noisy, while the effective training sample is short. In this setting—characterized by a small sample size and correlated regressors—ordinary least squares estimation can suffer from high variance and poor out-of-sample stability. Regularization is therefore essential to improve the bias–variance trade-off and stabilize forecasts.

3.10. Regularized Estimation

The forecasting environment considered in this paper combines a relatively small time series sample with predictors derived from high-dimensional textual representations. In such settings, regularized regression methods provide a natural framework for improving

the stability and out-of-sample performance of linear forecasting models. For this reason, we employ Ridge and Elastic Net estimators in the empirical analysis.

Three features of the forecasting problem motivate the use of shrinkage estimators. First, the effective sample size is limited. The dataset contains a relatively small number of Copom meetings, and the expanding-window forecasting design further restricts the number of observations available for estimation at each step. When the number of predictors is not negligible relative to the sample size, ordinary least squares can exhibit high sampling variability and unstable forecasts. Regularization mitigates this problem by penalizing coefficient magnitudes, reducing estimator variance and improving the bias–variance trade-off in small samples.

Second, the explanatory variables include factors derived from textual embeddings of central bank communication. Even after dimensionality reduction, these predictors may contain noise and redundancy, as some components may carry limited predictive signal for future policy decisions. Shrinkage methods help control overfitting by penalizing weak predictors. Ridge regression is particularly suitable when regressors are correlated—common with textual factors and dynamic regressors—while Elastic Net combines ridge-type shrinkage with an additional L_1 penalty that allows partial sparsity.

Third, the dependent variable, the change in the Selic policy rate, exhibits strong persistence and a large mass at zero due to monetary policy inertia. In this setting, incremental predictors may contain only weak information relative to the persistence component, making unrestricted estimation unstable. Shrinkage methods provide a disciplined way to incorporate additional predictors while controlling the influence of noisy signals. Taken together, these considerations make regularized regression well suited to the forecasting problem studied in this paper, allowing the model to incorporate information from central bank communication while maintaining stable out-of-sample performance in a small-sample environment.

Estimation is carried out using Ridge regression with an L_2 penalty applied to the slope coefficients. Operationally, after standardizing regressors using the training sample, we solve

$$\min_{\alpha,\rho,\beta} \frac{1}{n} \sum_{t \in \mathcal{D}_{\text{train}}} \left(y_{t+1} - \alpha - \rho\, y_t - \beta^\top f_t \right)^2 + \lambda \left(\rho^2 + \|\beta\|_2^2 \right), \tag{8}$$

where $\lambda \geq 0$ is selected using only training-sample information through time-series-appropriate validation or grid search. The intercept α is left unpenalized, while the penalty applies to both ρ and β, consistent with standard Ridge implementations in packages such as `scikit-learn`. Standardization ensures that the penalty treats coefficients associated with regressors on different scales in a comparable manner.

From a statistical perspective, this specification admits a natural interpretation in terms of sufficient statistics. Conditional on past policy decisions summarized by y_t, the vector of textual factors f_t is intended to act as a low-dimensional statistic that captures incremental information contained in the minutes about the conditional distribution of y_{t+1}. Ridge regularization plays a crucial role in this interpretation by shrinking weak or noisy components of this statistic toward zero. When the training data provide limited evidence that a given textual factor is informative, shrinkage prevents it from exerting undue influence on the forecast.

3.11. Robustness (PCA(k) + Elastic Net + lag(y_t))

As an additional robustness check, we estimate the same specification as the main model using Elastic Net regularization. The motivation is that, even after dimensionality reduction via PCA, textual factors may remain correlated and contain weak or unstable

signals in short samples. In this context, a method that combines shrinkage and sparsity can improve out-of-sample stability and mitigate overfitting.

The linear specification remains:

$$y_{t+1} = \alpha + \rho, y_t + \beta^\top f_t + \varepsilon_{t+1}, \tag{9}$$

but estimation is carried out through penalized least squares. Defining the regressor vector $X_t = (1, \ldots, y_t, \ldots, f_t^\top)$ and the parameter vector $\theta = (\alpha, \ldots, \rho, \ldots, \beta^\top)^\top$, the Elastic Net estimator solves:

$$\min_\theta \frac{1}{n} \sum_{t \in \mathcal{D}\text{train}} (yt + 1 - X_t\theta)^2 \lambda \Big[(1-\gamma)|\theta|_2^2 + \gamma|\theta|_1 \Big], \tag{10}$$

where $\lambda > 0$ controls the overall strength of regularization and $\gamma \in [0, 1]$ determines the balance between the L_2 and L_1 penalties. In particular, $\gamma = 0$ recovers Ridge regression, while $\gamma = 1$ corresponds to Lasso.

In the implementation, the intercept α is not penalized. The penalty applies to the coefficients associated with y_t and the textual factors f_t. Since the penalty depends on the scale of the regressors, all variables in X_t except the intercept are standardized on the training sample, and the same transformation is applied to the test observation within each window.

From a statistical perspective, Elastic Net enforces parsimony in the effective sufficient statistics extracted from the text. While PCA produces orthogonal factors summarizing information from high-dimensional embeddings, Elastic Net further restricts how much of this information is used for prediction. The L_2 component stabilizes estimation under collinearity and small samples, while the L_1 component allows the model to discard factors whose empirical covariance with y_{t+1} is weak or unstable.

This mechanism is particularly relevant when y_{t+1} has a large mass at zero. In such settings, many correlations between textual factors and outcomes may arise spuriously. Elastic Net mitigates this risk by concentrating predictive weight on a small subset of factors that consistently act as sufficient statistics for policy changes while shrinking diffuse or transitory signals. Consequently, the model remains close to the persistence benchmark in tranquil periods and reacts more strongly only when the textual signal is persistent and informative.

The hyperparameters (λ, γ) are selected exclusively on the training sample via time-series cross-validation (TimeSeriesSplit), preserving temporal ordering and avoiding information leakage. In each expanding window, the pair (λ, γ) that minimizes the validation error is chosen, after which the model is re-estimated on the full training sample to produce the out-of-sample forecast.

3.12. State-Space Robustness (Kalman/SARIMAX) with Exogenous Variables

As an econometric complement and a direct connection to state-space modeling and Kalman filtering, we estimate an autoregressive model with textual exogenous regressors in the SARIMAX framework. The one-step-ahead forecasting equation is:

$$y_{t+1} = c + \phi, y_t + \delta^\top f_t + u_{t+1}, \qquad u_{t+1} \sim \mathcal{N}(0, \sigma^2), \tag{11}$$

where $f_t \in \mathbb{R}^k$ denotes the textual factors available at time t, and $\delta \in \mathbb{R}^k$ is the corresponding loading vector.

The SARIMAX model admits a state-space representation, with an observation equation linking y_{t+1} to a latent autoregressive state and to exogenous textual factors, and a transition equation governing the evolution of the latent state. In this framework, the

Kalman filter provides recursive evaluation of the likelihood and optimal one-step-ahead projections under Gaussian assumptions, allowing estimation of $(c, \phi, \delta, \sigma^2)$ via maximum likelihood [33,44,45].

From the perspective of sufficient statistics, the state-space formulation provides a dynamic aggregation of information. The latent state summarizes past observations of y_t into a low-dimensional statistic sufficient for forecasting under the linear-Gaussian structure, while textual factors act as contemporaneous exogenous statistics that shift the conditional mean of y_{t+1}. In this sense, SARIMAX offers a structured benchmark combining persistence and textual information through probabilistic filtering rather than penalized least squares.

In each expanding window, the model is re-estimated using only the training sample, preserving temporal ordering, and produces an out-of-sample forecast $\hat{y}_{t+1|t}$. This procedure is directly comparable to the rolling estimation used for the penalized linear models but relies on likelihood-based inference and the state-space structure.

Estimation of SARIMAX models can be sensitive to initial conditions and local optima, particularly in short samples. To mitigate convergence issues, we adopt a robust optimization strategy with multiple solvers and increased iteration limits, recording convergence diagnostics in each window and reporting forecasts only when estimation converges successfully. In this way, the SARIMAX specification with exogenous textual factors serves as a robustness check, assessing whether the informational content of the text persists when dynamics are modeled within a classical state-space and Kalman filtering framework.

Although the empirical implementations differ, the Ridge, Elastic Net, and Kalman-based specifications can be interpreted within a unified framework in which forecasting relies on constructing low-dimensional sufficient statistics from noisy and highly correlated information sources.

In penalized linear models, regularization acts as an implicit information filter. Ridge regression stabilizes estimation by shrinking coefficients toward zero, effectively down-weighting weak or unstable correlations between regressors and y_{t+1}. Elastic Net extends this mechanism by combining shrinkage with sparsity: the L_1 component allows the model to discard entire directions in the factor space, producing a data-driven selection of textual components that act as relevant sufficient statistics.

The state-space model provides a probabilistic counterpart to this filtering logic. Under the linear-Gaussian structure, past realizations of y_t are compressed into a latent state that is sufficient for forecasting, and the Kalman filter recursively updates this state as new information arrives. Textual factors enter as exogenous inputs that shift the conditional mean and act as contemporaneous sufficient statistics.

Viewed jointly, Ridge and Elastic Net implement static filtering through regularization, whereas the Kalman filter performs dynamic filtering through recursive Bayesian updating. In all cases, the objective is to prevent the model from overreacting to noisy textual signals in a short-sample environment with high collinearity and a large mass of zero outcomes while still allowing persistent information in the text to influence forecasts.

The dependent variable—the change in the Selic target rate between consecutive Copom meetings—is discrete and highly concentrated at zero due to the frequency of rate-hold decisions. Despite this feature, we adopt a linear Gaussian forecasting framework motivated by practical considerations related to the forecasting objective and the structure of the available data.

First, the objective of the exercise is to generate out-of-sample forecasts of the magnitude of policy rate adjustments rather than to estimate the structural probability of discrete policy actions. A continuous framework allows the model to produce smooth predictions

that capture gradual changes in the expected policy stance, even when realized decisions are discrete, and facilitates evaluation using standard loss functions such as RMSE.

Second, the forecasting problem combines a short time series with a potentially high-dimensional set of textual predictors extracted from Copom minutes. Linear models with regularization methods such as Ridge and Elastic Net provide a stable and computationally tractable approach to handling this dimensionality while controlling estimation variance. Alternative discrete-choice frameworks, such as ordered logit or multinomial models, are less naturally compatible with high-dimensional shrinkage methods in small samples.

Finally, in forecasting applications it is common to approximate discrete outcomes using continuous predictive models when the objective is to minimize prediction error rather than to model the full decision process. In this sense, the Gaussian framework should be interpreted as a pragmatic approximation that prioritizes predictive stability and compatibility with regularized estimation.

3.13. Predictive Evaluation Design

Predictive performance is assessed using an expanding-window backtest designed to mimic real-time forecasting under information constraints. This protocol is particularly appropriate in environments with short samples and evolving information sets, as it avoids look-ahead bias and reflects the sequential nature of policy analysis.

Let t_0 denote the minimum training sample size. For each $t = t_0, \ldots, T-1$, the following steps are implemented:

1. define the training set as $\mathcal{D}_{\text{train}} = \{1, \ldots, t\}$;
2. estimate the forecasting model using only $\mathcal{D}_{\text{train}}$;
3. compute the textual factors f_t by applying a PCA transformation estimated exclusively on the training set to the test-period embedding;
4. generate the one-step-ahead forecast $\hat{y}_{t+1|t}$ and compare it with the realized outcome y_{t+1}.

This procedure yields $N = T - t_0 = 25$ genuine out-of-sample forecasts. By re-estimating all model components at each iteration, including dimensionality reduction and hyperparameters, the backtest evaluates not only the forecasting equation but the entire information-processing pipeline under realistic conditions.

3.14. Performance Metrics

Forecast accuracy is primarily evaluated using root mean squared error (RMSE) and mean absolute error (MAE):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=t_0}^{T-1}(y_{t+1} - \hat{y}_{t+1|t})^2}, \tag{12}$$

$$\text{MAE} = \frac{1}{N}\sum_{t=t_0}^{T-1}\left|y_{t+1} - \hat{y}_{t+1|t}\right|. \tag{13}$$

These metrics capture complementary aspects of forecast quality: RMSE penalizes large errors more heavily and is sensitive to tail events, while MAE provides a more robust measure under asymmetric or heavy-tailed forecast errors. Given the empirical distribution of y_{t+1}—characterized by a large mass at zero and occasional discrete jumps—reporting both metrics is informative.

As a directional diagnostic, we also compute sign accuracy:

$$\text{sign\_acc} = \frac{1}{N}\sum_{t=t_0}^{T-1}\mathbf{1}\{\text{sign}(y_{t+1}) = \text{sign}(\hat{y}_{t+1|t})\}.$$

Because both realizations and forecasts may be close to zero in magnitude, small numerical deviations may generate economically irrelevant sign disagreements. To address this issue, we additionally report a tolerance-based version using threshold τ (here, $\tau = 12.5$ basis points):

$$\text{sign}_\tau(x) = \begin{cases} 1, & x > \tau, \\ 0, & |x| \leq \tau, \\ -1, & x < -\tau, \end{cases} \tag{14}$$

and we compute the proportion of matches between $\text{sign}_\tau(y_{t+1})$ and $\text{sign}_\tau(\hat{y}_{t+1|t})$. This metric aligns directional evaluation with the discrete nature of monetary policy adjustments and avoids over-penalizing near-zero forecast errors.

3.15. Operational Choices and Experimental Hyperparameters

To enhance reproducibility and limit researcher degrees of freedom, the empirical exercise is conducted under a fixed and pre-specified set of operational choices:

- **Sample construction.** The sample includes Copom minutes starting from the 200th meeting, yielding a total of $T = 75$ observations.
- **Forecast horizon and protocol.** The forecasting horizon is one step ahead ($h = 1$), predicting y_{t+1} using information available at time t. Evaluation is conducted via an expanding-window backtest with minimum training size $t_0 = 50$, producing 25 out-of-sample forecasts.
- **Textual embeddings.** Each set of minutes is converted from PDF to text and partitioned into chunks. For each chunk, a dense embedding $e_{t,c} \in \mathbb{R}^{384}$ is computed. The meeting-level embedding is defined as the simple average across chunks,

$$e_t = \frac{1}{C_t} \sum_{c=1}^{C_t} e_{t,c}.$$

- **Textual factors.** Dimensionality reduction is performed via PCA, yielding $f_t \in \mathbb{R}^k$ with $k \in \{1, 2, 3, 4, 5\}$. In each backtest iteration, PCA is estimated exclusively on the training sample and then applied to the test observation (*train-only PCA*), ensuring strict separation between training and evaluation information.
- **Model specifications.** We compare simple benchmarks, a pure-text model (PCA(k) + OLS), the main regularized model with persistence (PCA(k) + Ridge + lag(y_t)), and robustness extensions using Elastic Net and SARIMAX with textual exogenous regressors.
- **Hyperparameter selection.** For regularized models, hyperparameters are selected using training data only. Elastic Net parameters are chosen via time-series cross-validation (TimeSeriesSplit). For the state-space model, estimation relies on maximum likelihood with explicit convergence checks in each window.
- **Evaluation metrics.** RMSE is the primary criterion for model comparison, with MAE reported as a complementary measure. Directional diagnostics include standard sign accuracy and the tolerance-based version with $\tau = 12.5$ bps.
- **Reproducibility.** All experiments are run with fixed random seed (`seed = 123`), and model configurations are logged using a standardized experimental protocol.

The preferred specification is defined ex ante as the model that minimizes out-of-sample RMSE, holding constant the backtesting protocol and all textual preprocessing steps. This design ensures that differences in predictive performance can be attributed to the forecasting models themselves rather than to variation in information sets or evaluation procedures.

As complementary evidence of robustness, the best-performing model is compared to the persistence benchmark using the Diebold–Mariano (DM) test, following [26], at horizon $h = 1$. The loss function is quadratic, $\ell_{t+1} = (y_{t+1} - \hat{y}_{t+1|t})^2$, and the loss differential is defined as

$$d_{t+1} = \ell_{t+1}^{(\text{model})} - \ell_{t+1}^{(\text{pers})}. \quad (15)$$

The test evaluates the null hypothesis $H_0 : \mathbb{E}[d_{t+1}] = 0$ against the alternative $H_1 : \mathbb{E}[d_{t+1}] \neq 0$.

Appendix A presents the computational details and a general diagram of the method.

4. Results

This section presents the out-of-sample forecasting results from the expanding-window backtest and evaluates the incremental predictive content of textual information from Copom minutes relative to standard benchmarks. The analysis focuses on one-step-ahead forecasts of changes in the Selic target rate, with performance assessed primarily through RMSE, complemented by MAE and directional accuracy measures.

Table 1 reports out-of-sample predictive performance under the expanding-window backtesting protocol, comparing benchmark time-series models to a range of text-based and text-augmented specifications. The evaluation focuses on one-step-ahead forecasts ($h = 1$) over $N = 25$ test observations, with RMSE in basis points as the primary performance criterion, complemented by MAE and directional accuracy measures. This comparison is designed to assess not only whether textual information extracted from Copom minutes improves forecast accuracy relative to standard benchmarks but also how such information interacts with short-run policy dynamics and regularization in a small-sample, high-noise environment.

Table 1. Out-of-sample predictive performance ($N = 25$ predictions, horizon $h = 1$). Main criterion: RMSE (bps). Relative RMSE and MAE are computed with respect to the persistence benchmark.

Model	k	RMSE	Rel. RMSE	MAE	Rel. MAE	N
Benchmarks						
RW (0 bps)	–	45.5522	2.15	31.0000	2.58	25
Persistence (y_t)	–	21.2132	1.00	**12.0000**	**1.00**	25
Pure Text (PCA + OLS)						
Text: PCA(k) + OLS	1	51.4314	2.43	43.3085	3.61	25
Text: PCA(k) + OLS	2	54.3865	2.56	45.5517	3.80	25
Text: PCA(k) + OLS	3	51.3364	2.42	40.9716	3.41	25
Text: PCA(k) + OLS	4	53.5048	2.52	43.4773	3.62	25
Text: PCA(k) + OLS	5	41.4287	1.95	31.8339	2.65	25
Text + dynamics (Ridge + lag)						
Text: PCA(k) + Ridge + lag(y_t)	1	21.1676	1.00	13.8544	1.15	25
Text: PCA(k) + Ridge + lag(y_t)	2	21.3903	1.01	13.3955	1.12	25
Text: PCA(k) + Ridge + lag(y_t)	3	21.4304	1.01	14.4513	1.20	25
Text: PCA(k) + Ridge + lag(y_t)	4	21.0814	0.99	14.2268	1.19	25
Text: PCA(k) + Ridge + lag(y_t)	**5**	**20.2987**	**0.96**	14.6901	1.22	25
SARIMAX with textual exogenous variables						
SARIMAX(1,0,0) + Text (PCA, exog)	1	21.1619	1.00	13.1023	1.09	25
SARIMAX(1,0,0) + Text (PCA, exog)	2	21.4363	1.01	14.5298	1.21	25
SARIMAX(1,0,0) + Text (PCA, exog)	3	20.9504	0.99	14.0230	1.17	25
SARIMAX(1,0,0) + Text (PCA, exog)	4	20.6029	0.97	13.9626	1.16	25
SARIMAX(1,0,0) + Text (PCA, exog)	5	20.9029	0.99	14.0057	1.17	25

Note: RMSE and MAE are in basis points (bps). Relative RMSE and MAE are computed with respect to the persistence benchmark. PCA is estimated only on the training sample in each expanding-window iteration. Bold values indicate the best performance in terms of RMSE and MAE.

The results are organized into three blocks: baseline macro-financial models, text-only specifications, and models combining macroeconomic and textual information. The persistence model provides a strong benchmark due to the high frequency of rate-hold decisions,

and regularized macro-financial models yield only modest improvements relative to this baseline. Models based exclusively on textual factors extracted from COPOM minutes generally perform worse than macro-financial benchmarks, particularly during periods of policy inertia, suggesting that textual signals are less informative when the policy rate remains unchanged. In contrast, hybrid models that combine macroeconomic predictors with textual factors deliver the strongest performance overall, indicating that central bank communication provides complementary information not captured by conventional variables. Notably, improvements in relative RMSE are concentrated during policy transition episodes, when forward-looking guidance in communication becomes more informative about the future policy path.

The persistence benchmark ($\hat{y}_{t+1|t} = y_t$) attains high directional accuracy (sign_acc $= 0.84$), reflecting the discrete and highly concentrated nature of the target variable, with substantial mass on rate-hold decisions ($y_{t+1} = 0$) and serially correlated sequences of tightening or easing. Economically, this pattern mirrors the well-known inertia of monetary policy: once the Monetary Policy Committee (Copom) reaches a level of the Selic rate considered appropriate, the policy stance is often maintained across several meetings, while the effects of previous adjustments propagate through the economy. In such environments, predicting no change becomes a strong baseline strategy, which explains the high directional performance of persistence. In contrast, the RW (0 bps) benchmark exhibits weak performance in terms of RMSE and MAE, even though it mechanically achieves a non-negligible sign accuracy when realizations cluster close to zero.

When textual information is incorporated in isolation, the text-only baseline (PCA(k) + OLS) underperforms the benchmarks, with higher RMSE and lower directional accuracy. This outcome is consistent with the challenges of relying exclusively on textual factors in short samples, where estimation uncertainty and the absence of explicit dynamics can dominate. From an economic perspective, it also suggests that Copom minutes rarely signal policy decisions in isolation. During prolonged maintenance phases, communication tends to reiterate previously stated assessments of inflation, activity, and risks, reinforcing the prevailing stance rather than providing immediate signals of policy adjustments.

The central result is that the text-augmented dynamic model with regularization—PCA(k) + Ridge + lag(y_t)—improves upon persistence in terms of the primary evaluation criterion (RMSE), with the strongest performance attained at $k = 5$ (RMSE $= 20.30$ bps). From an economic perspective, this finding is consistent with the interpretation that the minutes contain incremental information about both the likelihood and the magnitude of future policy changes, while short-run policy inertia, summarized by y_t, remains a key driver of predictability. In particular, textual signals appear to become most informative around turning points in the monetary policy cycle, when shifts in the narrative of the minutes reflect evolving assessments of inflation risks and the appropriate policy stance before they fully materialize in policy decisions.

The robustness exercises reinforce this interpretation. The Elastic Net specification with a lag delivers performance close to that of Ridge, indicating that the gains primarily reflect shrinkage and variance control rather than aggressive sparsity or selection of a small subset of factors. The SARIMAX specifications with textual exogenous variables also yield competitive RMSE in some configurations, supporting the view that the textual signal is compatible with a classical state-space representation estimated via maximum likelihood. At the same time, these results should be interpreted with caution given occasional convergence issues in short samples. Overall, the evidence suggests that textual information does not replace persistence but rather complements it by providing forward-looking signals about regime transitions in the monetary policy cycle.

Given that the target variable is discrete and exhibits substantial mass at zero, standard directional accuracy measures may be dominated by forecasts classified as maintenance decisions. To address this issue, in addition to the conventional sign_acc metric, we also report sign_acc_τ with $\tau = 12.5$ basis points, which treats predictions of small magnitude as neutral. Under this tolerance-based metric, the persistence benchmark continues to display strong performance, while models combining text and dynamics tend to achieve comparable values of sign_acc_τ. This suggests that the main contribution of textual information manifests primarily in improving the level accuracy of forecasts—as reflected in lower RMSE—rather than in simple directional classification, especially when the tolerance threshold effectively reframes the task as distinguishing between change and no-change decisions.

Given that the target variable is discrete and exhibits substantial mass at zero, standard directional accuracy measures may be dominated by forecasts that simply predict policy maintenance. To provide a more informative assessment, in addition to the conventional sign accuracy metric ($\tau = 0$), we report directional accuracy for a range of tolerance thresholds τ between 2.5 and 20 basis points. Under this tolerance-based classification, both predicted and realized changes within the interval $[-\tau, \tau]$ are treated as neutral decisions.

Table 2 shows that the persistence benchmark achieves very high directional accuracy across all tolerance levels. Economically, this reflects a key feature of monetary policy in Brazil: the Selic rate is often maintained for several consecutive meetings once the desired policy stance is reached. In such an environment, simply predicting no change is frequently correct, which makes persistence a particularly difficult benchmark to outperform in terms of direction alone.

Table 2. Sign accuracy for different tolerance levels (τ). Each column reports sign_acc_τ with tolerance τ (bps). N is the number of out-of-sample predictions.

Model	k	N	$\text{sign\_acc}_{\tau=0}$	$\text{sign\_acc}_{\tau=2.5}$	$\text{sign\_acc}_{\tau=5}$	$\text{sign\_acc}_{\tau=7.5}$	$\text{sign\_acc}_{\tau=10}$	$\text{sign\_acc}_{\tau=12.5}$	$\text{sign\_acc}_{\tau=15}$	$\text{sign\_acc}_{\tau=20}$
Benchmarks										
RW (0 bps)	–	25	0.440	0.440	0.440	0.440	0.440	0.440	0.440	0.440
Persistence (y_t)	–	25	0.840	0.840	0.840	0.840	0.840	0.840	0.840	0.840
Pure Text (PCA + OLS)										
Text: PCA(k) + OLS	1	25	0.320	0.360	0.360	0.320	0.280	0.280	0.280	0.280
Text: PCA(k) + OLS	2	25	0.240	0.280	0.280	0.280	0.280	0.240	0.200	0.160
Text: PCA(k) + OLS	3	25	0.360	0.320	0.280	0.360	0.280	0.320	0.360	0.320
Text: PCA(k) + OLS	4	25	0.320	0.240	0.240	0.280	0.240	0.280	0.320	0.320
Text: PCA(k) + OLS	5	25	0.400	0.440	0.440	0.400	0.400	0.400	0.480	0.480
Text + dynamics (Ridge + lag)										
Text: PCA(k) + Ridge + lag(y_t)	1	25	0.520	0.640	0.840	0.840	0.840	0.840	0.840	0.840
Text: PCA(k) + Ridge + lag(y_t)	2	25	0.520	0.800	0.840	0.840	0.840	0.840	0.840	0.840
Text: PCA(k) + Ridge + lag(y_t)	3	25	0.520	0.640	0.720	0.840	0.840	0.840	0.840	0.840
Text: PCA(k) + Ridge + lag(y_t)	4	25	0.520	0.640	0.760	0.800	0.840	0.840	0.840	0.840
Text: PCA(k) + Ridge + lag(y_t)	5	25	0.520	0.600	0.680	0.800	0.880	0.840	0.840	0.880
SARIMAX with textual exogenous variables										
SARIMAX(1,0,0) + Text (PCA, exog)	1	25	0.480	0.800	0.840	0.840	0.840	0.840	0.840	0.840
SARIMAX(1,0,0) + Text (PCA, exog)	2	25	0.520	0.760	0.760	0.760	0.800	0.840	0.840	0.840
SARIMAX(1,0,0) + Text (PCA, exog)	3	25	0.560	0.720	0.800	0.800	0.840	0.840	0.840	0.840
SARIMAX(1,0,0) + Text (PCA, exog)	4	25	0.520	0.600	0.800	0.800	0.800	0.840	0.840	0.840
SARIMAX(1,0,0) + Text (PCA, exog)	5	25	0.520	0.600	0.800	0.800	0.800	0.840	0.840	0.840

Note: sign_acc_τ compares the discretized sign of y_{t+1} and $\hat{y}_{t+1|t}$ using a tolerance τ (bps), where values in $[-\tau, \tau]$ are classified as 0. Rows are evaluated on the common out-of-sample dates.

Purely textual models based on PCA and OLS display relatively weak directional performance. This result suggests that textual information extracted from Copom minutes does not, by itself, reliably signal the direction of immediate policy changes. Instead, communication appears to convey richer information about the policy stance rather than explicit indications of imminent rate adjustments. In contrast, models that combine textual factors with policy dynamics—such as PCA(k) + Ridge + lag(y_t) and SARIMAX with textual exogenous variables—achieve substantially higher directional accuracy. As the tolerance threshold increases, these hybrid models rapidly converge to levels close to the persistence

benchmark. Economically, this indicates that once small fluctuations around zero are treated as neutral decisions, the models correctly capture the broad policy stance conveyed by central bank communication.

Overall, these results suggest that the primary contribution of textual information lies in refining forecasts of the magnitude and timing of policy adjustments rather than simply predicting their direction. While persistence captures the inertia inherent in monetary policy decisions, textual signals help identify changes in the policy environment that precede turning points in the interest-rate cycle. Consequently, the informational value of central bank communication becomes most apparent when forecasting the size of adjustments around policy transitions rather than during prolonged periods of rate stability.

To complement the numerical evaluation of forecast performance, Figures 1 and 2 provide a visual comparison between the persistence benchmark and the text-augmented model. These figures are useful for illustrating how differences in average performance metrics arise over time and for highlighting the specific episodes that account for improvements in RMSE.

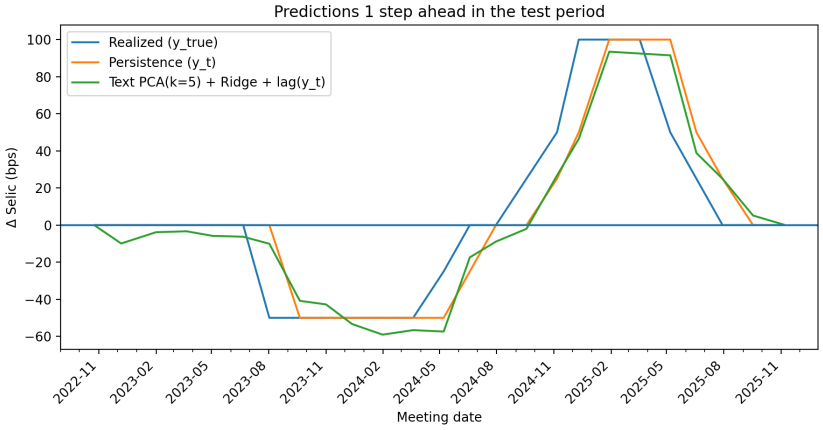

Figure 1. Predictions 1 step ahead in the test period. Comparison between the actual (y_{t+1}), the persistence benchmark ($\hat{y}_{t+1|t} = y_t$) and the model with text (PCA(5) + Ridge + lag(y_t)).

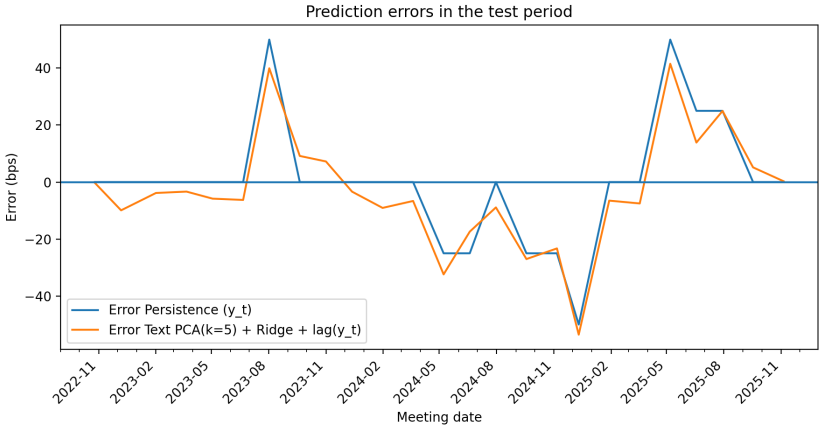

Figure 2. Prediction errors ($y_{t+1} - \hat{y}_{t+1|t}$) in the test period, persistence versus model with text (PCA(5) + Ridge + lag(y_t)).

Figure 1 provides a visual comparison between realized policy rate changes and the forecasts generated by the main models. The figure highlights the strong inertia of monetary policy decisions in Brazil: long stretches of unchanged policy rates are clearly visible, interspersed with relatively short tightening or easing cycles. In these plateau phases, the persistence benchmark closely tracks realized outcomes because predicting no

change is often correct. The model augmented with textual information exhibits a similar behavior during stable periods, but it tends to adjust more rapidly around turning points in the monetary policy cycle. This pattern suggests that shifts in the tone and content of Copom minutes may contain early signals of evolving policy assessments, allowing the text-based model to react sooner when the macroeconomic environment begins to change.

Figure 2 complements this interpretation by showing the temporal distribution of forecast errors. The figure indicates that most of the quadratic loss is driven by a relatively small number of meetings associated with policy transitions. During these episodes, persistence forecasts can generate large errors because they extrapolate the previous policy decision even when the central bank is about to change direction. In contrast, models incorporating textual information tend to produce smaller errors in these periods, reflecting their ability to capture forward-looking signals embedded in central bank communication. Economically, this result is consistent with the role of Copom minutes as a vehicle for communicating evolving assessments of inflation risks, economic activity, and the appropriate policy stance. As a result, textual information appears to be particularly informative in environments where the policy stance is shifting while providing limited incremental value during prolonged periods of rate maintenance.

4.1. Statistical Comparison (Diebold–Mariano)

As a robustness check, we apply the Diebold–Mariano (DM) test [26] to compare the model with the lowest RMSE, namely PCA($k = 5$) + Ridge + lag(y_t), against the persistence benchmark under quadratic loss,

$$\ell_{t+1} = (y_{t+1} - \hat{y}_{t+1|t})^2,$$

at the one-step-ahead horizon ($h = 1$). Table 3 reports the average losses over the test period and the mean loss differential $\bar{d} = \overline{\ell_{\text{model}} - \ell_{\text{persist}}}$.

Table 3. Diebold–Mariano (DM) test during the test period ($h = 1$, quadratic loss). We report both the original DM statistic and the [46] corrected statistic $DM^* = DM \times 0.9798$, where $0.9798 = \sqrt{(N-1)/N}$ with $N = 25$ (two-sided t_{N-1} approximation). * Denotes significante at 5% level.

Model	mean_loss	$\bar{d}$	*DM*	*p*-Value	DM*	*p*-Value *
Persistence (y_t)	450.00	0.00	–	–	–	–
Text: PCA(k = 5) + Ridge + lag(y_t)	412.04	**−37.96**	−0.65	0.52	−0.64	0.53

The text-based model exhibits a lower mean loss (mean_loss = 412.04) than the persistence benchmark (mean_loss = 450.00), implying $\bar{d} = -37.96$. Thus, on average, the model produces smaller squared forecast errors than the benchmark.

Despite this reduction in average loss, the original DM test does not reject the null hypothesis of equal predictive accuracy (test statistic $DM = -0.65$, $p = 0.52$). Given the relatively small test sample ($N = 25$), we also report the small-sample modified statistic proposed by Harvey et al. [46]. This correction rescales the DM statistic by the factor $\sqrt{(N-1)/N}$, producing $DM^* = -0.64$ with a corresponding p-value of 0.53. The modified test leads to the same qualitative conclusion as the original statistic.

Overall, both the standard and modified DM statistics indicate that the null hypothesis of equal predictive accuracy cannot be rejected at conventional significance levels. This result should be interpreted in light of two features of the empirical setting. First, the test sample is relatively small, which limits the statistical power of forecast comparison tests. Second, quadratic loss can be influenced by a small number of meetings with unusually large forecast errors, increasing the variability of the loss differential. Therefore, the DM results are best viewed as complementary evidence: while the text-based specification

achieves lower RMSE and lower average loss than the persistence benchmark, the statistical strength of this improvement remains limited given the sample size and the inherent variability of the loss differential.

Note that the losses reported in Table 3 correspond to averages of $(y_{t+1} - \hat{y}_{t+1|t})^2$ over the test period, and that negative values of the loss differential indicate lower average loss relative to the persistence benchmark.

4.2. Regime-Based Diagnostics and Economic Interpretation

A salient regularity in the results is that the incremental predictive value of textual information arises primarily in meetings associated with policy rate changes ($y_{t+1} \neq 0$). Table 4 decomposes out-of-sample performance into episodes of rate maintenance ($y_{t+1} = 0$) and rate adjustments ($y_{t+1} \neq 0$). During maintenance periods, the persistence benchmark remains highly competitive, reflecting both the strong inertia of monetary policy decisions and the substantial mass at zero in the target variable. In contrast, in episodes of policy change, the text-augmented model (PCA($k = 5$) + Ridge + lag) tends to generate smaller errors in magnitude, contributing disproportionately to the improvement observed in the overall RMSE (Table 1).

Table 4. Conditional performance in maintenance meetings ($y_{t+1} = 0$) and change meetings ($y_{t+1} \neq 0$).

Model	$N_{y=0}$	$\text{RMSE}_{y=0}$	$\text{MAE}_{y=0}$	$N_{y\neq0}$	$\text{RMSE}_{y\neq0}$	$\text{MAE}_{y\neq0}$
Persistence (y_t)	11	10.66	4.55	14	26.73	17.86
Text:PCA($k = 5$) + Ridge + lag(y_t)	11	10.57	7.84	14	25.46	20.08
RW (0 bps)	11	0.00	0.00	14	60.87	55.36

Note: Errors in basis points (bps). The test period contains $N = 25$ predictions, with $N_{y=0} = 11$ maintenance meetings and $N_{y\neq0} = 14$ meetings with rate changes. Since the RW benchmark (0 bps) always predicts maintenance, it is mechanically perfect in the $y_{t+1} = 0$ regime (zero RMSE and MAE) but performs poorly when $y_{t+1} \neq 0$. Among the models that incorporate dynamics, persistence remains competitive in maintenance, while the text-augmented model reduces the error in the rate change regime (lower $\text{RMSE}_{y\neq0}$).

Table 4 highlights a clear regime dependence in forecasting performance. During maintenance meetings, persistence-based forecasts perform strongly because the most common outcome is simply no change in the policy rate. In contrast, when policy adjustments occur, the specification incorporating textual factors achieves lower RMSE than persistence, indicating that information embedded in Copom communication helps anticipate policy moves.

This pattern is also visible in Figure 1. The persistence benchmark closely tracks extended periods of policy stability, whereas the model augmented with textual information adjusts more rapidly around turning points in the monetary policy cycle. Figure 2 further shows that a small number of meetings associated with policy transitions generate large forecast errors and therefore dominate the quadratic loss. Improvements during these episodes therefore have a disproportionate impact on RMSE performance.

The regime-dependent behavior of the forecasts reflects an important institutional feature of monetary policy in Brazil. The Central Bank of Brazil (BCB) typically conducts policy in cycles characterized by sequences of rate adjustments followed by prolonged periods of stability. Once the Monetary Policy Committee (Copom) reaches a level of the Selic rate considered sufficiently restrictive or accommodative, the policy rate is often held constant across multiple meetings while the committee evaluates the lagged effects of previous decisions on inflation and economic activity.

Recent policy developments illustrate this pattern. Between late 2024 and mid-2025, the BCB implemented a tightening cycle that raised the Selic rate to 15% per year. After this sequence of increases, the Copom interrupted the tightening cycle and maintained

the policy rate at that level for several meetings while monitoring inflation dynamics and economic activity [47–49]. In such environments, persistence forecasts become particularly difficult to outperform because the empirical distribution of outcomes is dominated by zero changes.

At the same time, episodes surrounding policy turning points create a different informational environment. When inflation risks, financial conditions, or the macroeconomic outlook begin to change, Copom communication often reflects evolving assessments of the appropriate policy stance. Shifts in the tone and content of the minutes may therefore precede actual rate adjustments.

The empirical results are consistent with this mechanism. Forecasting improvements from incorporating textual information are concentrated around policy transition episodes rather than during prolonged rate-hold periods. Economically, this suggests that central bank communication provides forward-looking signals about changes in the policy stance while naturally adding little incremental predictive value when the committee intentionally maintains a stable policy rate.

Overall, the results illustrate how the informational value of central bank communication depends on the policy regime. Persistence remains a strong benchmark during maintenance phases, reflecting deliberate policy inertia, whereas textual signals become most informative when the macroeconomic environment is evolving and policy adjustments become more likely.

4.3. Discussion

The evidence suggests that textual information from Copom minutes adds predictive value primarily when combined with short-run dynamics and regularization. Text alone performs worse than standard time-series benchmarks, indicating that textual factors cannot replace the information embedded in lagged policy decisions. However, when textual embeddings are incorporated together with the lagged policy rate and estimated with shrinkage methods such as Ridge, forecasting performance improves relative to persistence, particularly in terms of RMSE while preserving comparable directional accuracy. This pattern indicates that the predictive content of central bank communication becomes useful only once it is embedded in a framework that respects both the temporal structure of policy decisions and the need for statistical regularization.

From the perspective of sufficient statistics, these results imply that textual embeddings extracted from the minutes are not sufficient summaries of future policy decisions on their own. Instead, sufficiency emerges only conditionally when textual information is combined with the lagged policy rate. The lagged rate captures the slow-moving component of the monetary policy reaction function, reflecting the well-known inertia of policy decisions, while textual embeddings provide incremental information about potential deviations from this baseline. In particular, changes in the tone or emphasis of the minutes may signal reassessments of inflation risks, shifts in the balance of probabilities, or forward-looking guidance about the policy stance.

Regularization plays a central role in making this joint information structure empirically viable. Even after dimensionality reduction through PCA, textual factors remain noisy and weakly identified in small samples. Penalized estimators such as Ridge act as a statistical filter that stabilizes estimation and prevents noisy textual variation from overwhelming the persistent signal contained in the lagged policy rate. The weaker performance of PCA + OLS relative to PCA + Ridge illustrates that without regularization, textual variation is more likely to reflect estimation noise rather than genuine policy signals.

The regime decomposition further reinforces this interpretation. The gains from incorporating textual information are concentrated in meetings associated with policy

changes, where models augmented with textual factors achieve lower mean squared error. In contrast, persistence-based forecasts remain highly competitive during rate-maintenance episodes, when the policy rate is deliberately held constant for several meetings. This state-dependent performance is consistent with the economic role of central bank communication: during extended plateaus, the minutes largely reaffirm an unchanged policy stance, whereas during transition periods, they tend to reflect evolving assessments of macroeconomic conditions and risks.

This regime-specific behavior is also consistent with the visual evidence from the time-series plots of predictions and errors, where the largest improvements from text occur around turning points in the monetary policy cycle. From a filtering perspective, textual embeddings can therefore be interpreted as noisy signals about changes in an underlying policy state. Their informational value increases when that state is shifting, while it naturally diminishes during periods of policy stability.

Monetary policy typically exhibits substantial inertia. Central banks tend to adjust policy rates gradually and maintain the policy stance across several meetings once an appropriate level has been reached. As a result, the distribution of policy rate changes contains a large mass at zero, reflecting extended periods of unchanged rates. In such environments, persistence-based forecasts constitute a strong benchmark because the most likely outcome at any given meeting is often no change in the policy rate. Statistically, the high frequency of zero outcomes strengthens the predictive power of simple autoregressive structures and limits the scope for additional predictors to improve forecast accuracy.

This feature has important implications for models incorporating textual information from central bank communication. During prolonged periods of policy stability, the informational content of official communication regarding imminent policy adjustments is often limited. The minutes frequently reiterate previous assessments of inflation dynamics, economic activity, and risks while reaffirming the current policy stance. In these circumstances, textual signals tend to confirm the prevailing policy position rather than provide new information about future policy moves, implying a small incremental predictive contribution relative to persistence.

The forecasting environment therefore exhibits a form of regime dependence. Persistence-based models tend to perform best during maintenance phases, whereas textual communication becomes more informative when the policy stance is evolving. Episodes of tightening or easing are often accompanied by shifts in the central bank's narrative about inflation risks, macroeconomic conditions, or policy trade-offs. These changes may appear in the language of the minutes before they are fully reflected in policy decisions, allowing text-based predictors to capture early signals of upcoming adjustments. Consistent with this interpretation, the empirical results show that improvements from including textual factors are concentrated around policy transition episodes rather than during prolonged rate-hold periods. This suggests that textual communication is most informative about the timing and direction of policy shifts while naturally providing limited additional predictive value when the policy rate remains stable. The findings indicate that the predictive value of central bank communication varies across the monetary policy cycle. Textual information is most informative when the policy stance is shifting, as forward-looking signals embedded in communication can help anticipate upcoming rate adjustments. During periods of strong policy inertia, however, persistence-based forecasts remain difficult to outperform because policy decisions tend to remain unchanged across meetings.

Overall, the evidence suggests that textual information provides incremental forecasting power only when combined with dynamic structure and regularization. Rather than replacing persistence, text refines it by acting as a conditionally informative signal about regime transitions, consistent with a filtering perspective in which embeddings become

useful only after appropriate shrinkage and in conjunction with the evolving dynamics of policy decisions.

5. Conclusions

This study examined whether the textual content of Copom minutes contains predictive information about subsequent changes in the Selic policy rate. The results show that text alone is not sufficient to outperform simple benchmarks that exploit the strong persistence and discreteness of monetary policy decisions. Models based exclusively on textual factors tend to be unstable and fail to capture the dominant temporal dependence embedded in the policy cycle.

When textual information is combined with short-run dynamics and appropriate regularization, however, it contributes incremental predictive value. In particular, text-augmented models better anticipate policy adjustments and reduce forecast errors around turning points of the monetary policy cycle. This indicates that the informational content of the minutes is complementary rather than substitutive: it refines expectations about future decisions but does not replace the inertia embedded in past policy actions.

A central regularity is that the predictive gains from text are concentrated in episodes of rate changes. During periods of policy maintenance, persistence-based benchmarks remain highly competitive, reflecting the prevalence of unchanged decisions and the gradual nature of monetary policy. In contrast, when policy shifts occur, the language of the minutes appears to convey signals about evolving risks and forward-looking assessments that help anticipate both the direction and magnitude of future moves.

Although the textual factors are primarily used for prediction rather than structural interpretation, it is informative to examine the Copom meetings associated with extreme realizations of the principal components extracted from the minutes. These episodes typically coincide with major turning points in Brazil's monetary policy cycle, when central bank communication becomes richer in forward-looking signals. Several large factor realizations occur during periods of monetary tightening associated with elevated inflationary pressures. Meetings during the 2021–2022 tightening cycle, for example, correspond to strong positive realizations of the first textual factor. During this period, Copom repeatedly emphasized the deterioration of the inflation outlook and the need for rapid monetary policy normalization, highlighting persistent inflationary pressures from commodity prices, food and energy shocks, and exchange rate pass-through [50].

Another set of extreme realizations is associated with the later stages of the tightening cycle, when the Selic rate approached highly restrictive levels. Minutes from mid-2025 emphasize the resilience of domestic economic activity, elevated inflation expectations, and the need to maintain a contractionary stance for a prolonged period [47–49]. These discussions highlight concerns about the balance of risks and appear in the textual factors as large deviations from the average tone of the minutes.

Meetings associated with extreme negative textual factor realizations typically occur during transitions toward monetary easing, when Copom communication shifts toward discussing the moderation of inflation, the cumulative effects of previous tightening, and uncertainties surrounding the pace of policy normalization. In these episodes, the minutes often emphasize external risks, fiscal developments, and the lagged transmission of monetary policy. More broadly, extreme values of the textual factors tend to coincide with periods of elevated macroeconomic uncertainty and active changes in the policy stance, reinforcing the empirical finding that textual information from Copom minutes is most informative around turning points of the monetary policy cycle, when communication conveys stronger signals about the future path of interest rates.

The results support the view that central bank communication contains economically meaningful information, but that its value materializes only within models that respect the underlying dynamics of policy decisions and control for overfitting. Textual signals enhance forecasts primarily by sharpening responses at moments of adjustment rather than by improving performance in tranquil periods. These findings underscore both the potential and the limits of textual analysis in monetary policy applications and highlight the importance of integrating modern text representations with disciplined dynamic modeling frameworks.

Beyond statistical forecast evaluation, the results have direct implications for financial market participants exposed to monetary policy risk. In Brazil, expectations about the path of the Selic rate play a central role in the pricing and hedging of fixed-income securities, interest rate swaps, and derivatives linked to short-term rates. In particular, the Brazilian exchange (B3) lists derivatives whose payoffs are directly linked to Copom decisions, commonly referred to as Copom options, which allow market participants to hedge or speculate on the outcome of upcoming monetary policy meetings.

In this context, the forecasting framework developed in this paper can be interpreted as a tool for extracting incremental information from central bank communication that may complement market-based expectations. Because Copom minutes are released between policy meetings and often contain qualitative assessments of inflation dynamics, economic activity, and the balance of risks, text-based models provide a systematic way to translate this information into quantitative signals about the likelihood and magnitude of future policy adjustments. Our results suggest that such signals are most informative during periods in which the policy cycle is actively changing, episodes of monetary tightening or easing, precisely when uncertainty about future policy decisions tends to be higher and the pricing of policy-sensitive instruments becomes more volatile.

From a risk management perspective, improved forecasts of policy adjustments can enhance the calibration of scenario analyses and stress-testing exercises for interest rate exposures. For example, financial institutions managing fixed-income portfolios may use the forecasts to update the distribution of potential policy outcomes ahead of Copom meetings, improving hedging strategies in DI futures or options on policy decisions. Similarly, traders in Copom-linked derivatives may use signals extracted from the minutes as an additional input alongside market-implied expectations derived from the term structure of interest rates.

The study has several natural limitations. First, the sample is short and the target variable is discrete, which increases the difficulty of outperforming mechanical benchmarks and reduces the power of formal tests. In line with the concentration of observations at zero, purely autoregressive models naturally perform well in terms of RMSE by producing forecasts close to zero most of the time, even if they fail to capture rare but economically relevant rate changes. Second, there is an operational timing issue: the minutes are published after the meeting, so the results should be interpreted as evidence about the informational content of the text rather than as a strictly real-time forecasting strategy for the subsequent decision (a natural extension is to focus on documents available immediately after the decision (e.g., the statement) or to redefine the forecasting horizon to reflect the publication calendar of the minutes). Third, the use of embeddings combined with PCA imposes an aggregated representation of the text that may not fully exploit document structure, such as sections, tone, topics, or emphasis.

Several extensions are left for future research. Although the Gaussian forecasting framework provides a practical and stable approach in the present small-sample setting, several alternative modeling strategies could be explored in future research. Because Selic rate decisions are discrete and exhibit a substantial mass at zero, one natural extension

would be to consider ordered or multinomial specifications that explicitly model the probability of tightening, easing, or maintaining the policy rate. Another possibility would be to adopt hurdle-type or two-stage models that separately capture the probability of a rate change and the magnitude of the adjustment conditional on a change occurring. Classification-based approaches based on machine learning methods could also be used to focus directly on predicting the direction of policy moves. Exploring these alternative frameworks could provide additional insights into how textual information from central bank communication affects different dimensions of monetary policy predictability.

Another avenue is to incorporate additional textual sources from central bank communication at different horizons, such as the post-meeting statement, press conferences, or the Inflation Report. Measures of tone (e.g., hawkish versus dovish) could also be included as complements to embeddings. Further robustness checks could explore alternative evaluation schemes (e.g., rolling windows) and loss functions less sensitive to a small number of outliers. Finally, expanding the sample and investigating subperiods—particularly in light of changes in the monetary policy framework and communication style—would help assess the stability and generality of the results.

In summary, the evidence suggests that textual information from Copom minutes contains incremental predictive signal for Selic rate changes, but this value materializes primarily when text is combined with dynamic structure and regularization and is most pronounced during episodes of policy adjustment. At the same time, simple benchmarks remain difficult to outperform during rate-maintenance periods, and formal statistical evidence is constrained by sample size. Even so, the exercise demonstrates the feasibility of integrating modern textual representation techniques with rigorous out-of-sample evaluation protocols in applied monetary policy analysis.

Author Contributions: Conceptualization, I.B.d.A.D. and M.P.L.; methodology, I.B.d.A.D. and M.P.L.; validation, I.B.d.A.D. and M.P.L.; formal analysis, I.B.d.A.D. and M.P.L.; investigation, I.B.d.A.D. and M.P.L.; resources, M.P.L.; data curation, I.B.d.A.D.; writing—original draft preparation, I.B.d.A.D. and M.P.L.; writing—review and editing, I.B.d.A.D. and M.P.L.; visualization, I.B.d.A.D.; supervision, M.P.L.; funding acquisition, M.P.L. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge funding from Capes (Finance Code 001), CNPq (310646/2021-9) and FAPESP (2023/02538-0).

Data Availability Statement: Data from Central Bank of Brasil.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Computational Implementation

All code was implemented in Python (3.9) using a Jupyter Notebook v6 and executed in a local environment. The project is organized into standardized directories that store raw data, PDFs of the Copom minutes, and time series from the Central Bank of Brazil. These directories also contain the final panel used in the estimations, including the target variable and the embeddings aggregated at the meeting level, as well as tables and figures generated automatically.

The implementation is structured in stages covering project setup and configuration; collection and cleaning of the Selic target rate series and construction of the target variable; acquisition, extraction, and textual preprocessing of the minutes; generation and aggregation of embeddings by meeting; estimation and out-of-sample predictive evaluation using baseline models, text-based models, and robustness extensions; and the production of diagnostic tables and figures. At each stage, intermediate and final outputs are saved to disk, allowing for full auditing and replication of the results. The consolidated performance table

across models is stored, while additional diagnostics—by regime, error decomposition, and test-period plots—are saved in a separate directory.

Reproducibility is ensured through the use of a fixed pseudo-random seed, explicit storage of experimental configurations, and a standardized evaluation protocol. In particular, all transformations potentially subject to temporal leakage—most notably the estimation of the PCA used to construct textual factors and the selection of hyperparameters in regularized models—are performed exclusively using training data within each iteration of the expanding-window backtest. As a result, out-of-sample forecasts replicate an operational setting in which, at each meeting, only information available up to that point is used to update parameters and generate $\hat{y}_{t+1|t}$.

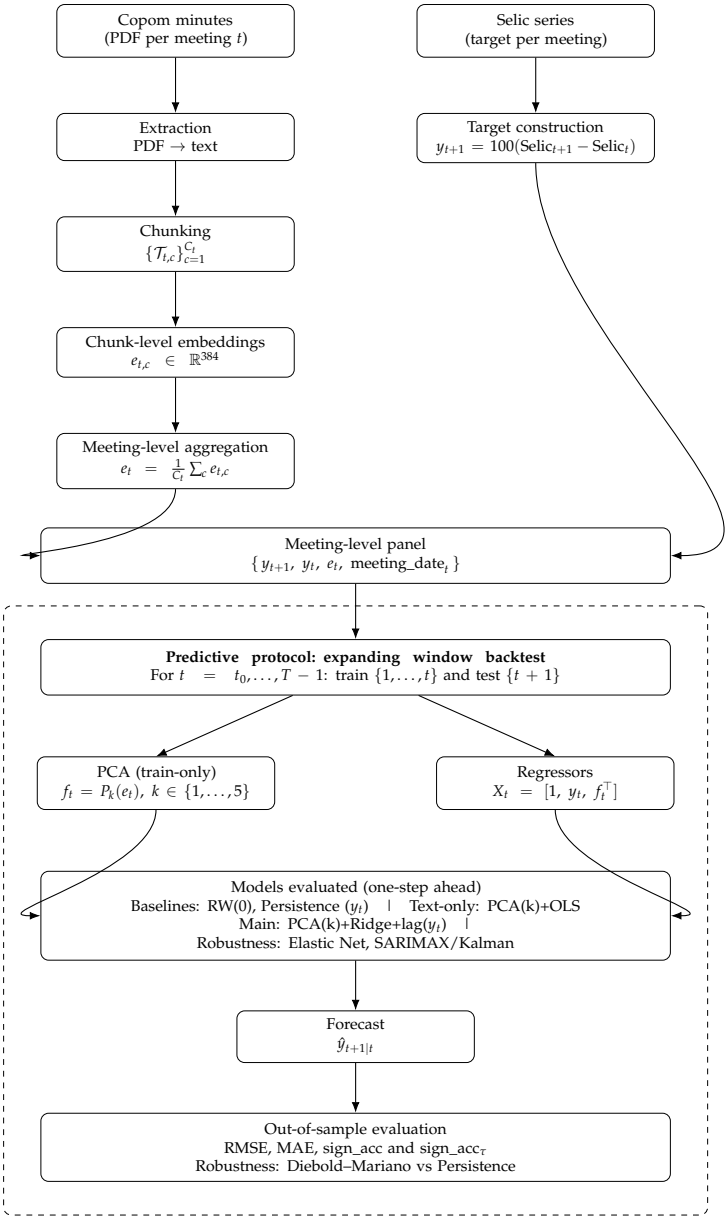

Figure A1. Empirical workflow: Copom minutes → embeddings → textual factors → models → expanding-window backtest and out-of-sample metrics.

References

1. Blinder, A.S.; Ehrmann, M.; Fratzscher, M.; De Haan, J.; Jansen, D.J. Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence. *J. Econ. Lit.* **2008**, *46*, 910–945. [CrossRef]

2. Carvalho, C.; Cordeiro, F.; Vargas, J. Just Words? A Quantitative Analysis of the Communication of the Central Bank of Brazil. *Rev. Bras. Econ.* **2013**, *67*, 443–455. [CrossRef]
3. Woodford, M. *Interest and Prices: Foundations of a Theory of Monetary Policy*; Princeton University Press: Princeton, NJ, USA, 2003.
4. Gürkaynak, R.S.; Sack, B.; Swanson, E.T. Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements. *Int. J. Cent. Bank.* **2005**, *1*, 55–93.
5. Campbell, J.R.; Evans, C.L.; Fisher, J.D.; Justiniano, A. Macroeconomic Effects of Federal Reserve Forward Guidance. *Brookings Pap. Econ. Act.* **2012**, *43*, 1–80. [CrossRef]
6. Nakamura, E.; Steinsson, J. High-Frequency Identification of Monetary Non-Neutrality: The Information Effect*. *Q. J. Econ.* **2018**, *133*, 1283–1330. [CrossRef]
7. Tadle, R.C. FOMC minutes sentiments and their impact on financial markets. *J. Econ. Bus.* **2022**, *118*, 106021. [CrossRef]
8. Gentzkow, M.; Kelly, B.; Taddy, M. Text as Data. *J. Econ. Lit.* **2019**, *57*, 535–574. [CrossRef]
9. Hansen, S.; McMahon, M. Shocking language: Understanding the macroeconomic effects of central bank communication. *J. Int. Econ.* **2016**, *99*, S114–S133. [CrossRef]
10. Rosa, C. The Impact of Federal Reserve Monetary Policy Announcements on Asset Prices: A Review. *Riv. Internazionale Sci. Soc.* **2019**, *127*, 3–24.
11. Usta, A.; Sert, M.F. Sentiment matters: Financial impact of tonal characteristics of FOMC communication. *Borsa Istanb. Rev.* **2025**, *26*, 100768. [CrossRef]
12. de Andrade Alves, C.R.; Joseph Abraham, K.; Poletti Laurini, M. Can Brazilian Central Bank communication help to predict the yield curve? *J. Forecast.* **2023**, *42*, 1429–1444. [CrossRef]
13. Alves, C.R.A.; Laurini, M.P. The effects of Brazilian Central Bank communication on the yield curve. *Macroecon. Dyn.* **2025**, *29*, e83. [CrossRef]
14. Cabral, R.; Guimarães, B. O Comunicado do Banco Central e a Ata do Copom São Redundantes? Uma Análise do Efeito dos Comunicados do Copom Sobre as Taxas de Juros. *Rev. Bras. Econ.* **2015**, *69*, 355–372. [CrossRef]
15. Chague, F.; De-Losso, R.; Giovannetti, B.; Manoel, P. Central Bank Communication Affects the Term Structure of Interest Rates. *Rev. Bras. Econ.* **2015**, *69*, 147–162. [CrossRef]
16. Ramos, F.S.; Portugal, M.S. O Poder da Comunicação do Banco Central: Avaliando o impacto sobre juros, bolsa, câmbio e expectativa de inflação. *Planej. Políticas Públicas* **2016**, *47*, 219–249.
17. Ash, E.; Hansen, S. Text Algorithms in Economics. *Annu. Rev. Econ.* **2023**, *15*, 659–688. [CrossRef]
18. Aruoba, S.B.; Drechsel, T. Identifying Monetary Policy Shocks: A Natural Language Approach. In *NBER Working Paper*; No. 32417; National Bureau of Economic Research: Cambridge, MA, USA, 2024. [CrossRef]
19. Baumgärtner, M.; Zahner, J. Whatever it takes to understand a central banker—Embedding their words using neural networks. *J. Int. Econ.* **2025**, *157*, 104101. [CrossRef]
20. Silva, T.C.; Moriya, K.; Veyrune, R.M. *From Text to Quantified Insights: A Large-Scale LLM Analysis of Central Bank Communication*; Technical Report WP/25/109; International Monetary Fund: Washington, DC, USA, 2025.
21. Lucchetti, A.H.; Cajueiro, D.O. A Multi-Dimensional Sentiment Index from Brazilian Central Bank Communication. *SSRN* **2025**. [CrossRef]
22. Moreno-Pérez, C.; Minozzo, M. The minutes of the Central Bank of Brazil and the real economy. *J. Int. Money Financ.* **2024**, *147*, 103133. [CrossRef]
23. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55–67. [CrossRef]
24. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B* **1996**, *58*, 267–288. [CrossRef]
25. Zou, H.; Hastie, T. Regularization and Variable Selection via the Elastic Net. *J. R. Stat. Soc. Ser. B* **2005**, *67*, 301–320. [CrossRef]
26. Diebold, F.X.; Mariano, R.S. Comparing Predictive Accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263. [CrossRef]
27. Ferreira, L.N.; Garzeri, C.; Guillen, D.; Lima, A.; Monteiro, V. *The Not So Quiet Revolution: Signal and Noise in Central Bank Communication*; Working Paper Series 635; Banco Central do Brasil: Brasília, Brazil, 2025.
28. Araujo, G.S.; Gaglianone, W.P. *Machine Learning Methods for Inflation Forecasting in Brazil: New Versus Old*; Working Paper Series 561; Banco Central do Brasil: Brasília, Brazil, 2022.
29. Stock, J.H.; Watson, M.W. Forecasting Using Principal Components from a Large Number of Predictors. *J. Am. Stat. Assoc.* **2002**, *97*, 1167–1179. [CrossRef]
30. Hastie, T.; Tibshirani, R.; Wainwright, M. *Statistical Learning with Sparsity: The Lasso and Generalizations*; CRC Press: Boca Raton, FL, USA, 2015.
31. Masini, R.P.; Medeiros, M.C.; Mendes, E.F. Machine learning advances for time series forecasting. *J. Econ. Surv.* **2023**, *37*, 76–111. [CrossRef]
32. Giannone, D.; Lenza, M.; Primiceri, G.E. Economic Predictions With Big Data: The Illusion of Sparsity. *Econometrica* **2021**, *89*, 2409–2437. [CrossRef]

33. Durbin, J.; Koopman, S.J. *Time Series Analysis by State Space Methods* ; Oxford University Press: Oxford, UK, 2012.
34. West, M.; Harrison, J. *Bayesian Forecasting and Dynamic Models*; Springer: Berlin/Heidelberg, Germany, 1997.
35. Fan, J.; Xue, L.; Yao, J. Sufficient forecasting using factor models. *J. Econom.* **2017**, *201*, 292–306. [CrossRef]
36. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
37. Tetlock, P.C. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *J. Financ.* **2007**, *62*, 1139–1168. [CrossRef]
38. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*; Burstein, J., Doran, C., Solorio, T., Eds.; Association for Computational Linguistics: New York, NY, USA , 2019; Volume 1 (Long and Short Papers), pp. 4171–4186. [CrossRef]
39. Ramshaw, L.A.; Marcus, M.P. Text Chunking using Transformation-Based Learning. In Proceedings of the Third Workshop on Very Large Corpora, Cambridge, MA, USA, 30 June 1995.
40. Timmermann, A. Forecast Combinations. In *Handbook of Economic Forecasting*, 1st ed.; Elliott, G., Granger, C., Timmermann, A., Eds.; Elsevier: Amsterdam, The Netherlands, 2006; Volume 1, Chapter 4, pp. 135–196.
41. Roberts, M.E.; Stewart, B.M.; Tingley, D.; Lucas, C.; Leder-Luis, J.; Gadarian, S.K.; Albertson, B.; Rand, D.G. Structural Topic Models for Open-Ended Survey Responses. *Am. J. Political Sci.* **2014**, *58*, 1064–1082. [CrossRef]
42. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. [CrossRef]
43. Wang, W.; Wei, F.; Dong, L.; Bao, H.; Yang, N.; Zhou, M. MINILM: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. In Proceedings of the 34th International Conference on Neural Information Processing Systems, NIPS '20, Red Hook, NY, USA, 6–12 December 2020.
44. Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng.* **1960**, *82*, 35–45. [CrossRef]
45. Alharbi, F.R.; Csala, D. A seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) forecasting model-based time series approach. *Inventions* **2022**, *7*, 94. [CrossRef]
46. Harvey, D.I.; Leybourne, S.J.; Newbold, P. Testing the Equality of Prediction Mean Squared Errors. *Int. J. Forecast.* **1997**, *13*, 281–291. [CrossRef]
47. Central Bank of Brazil. Minutes of the Monetary Policy Committee (Copom) Meeting, June 2025. Available online: https://www.bcb.gov.br/en/pressdetail/2617/nota (accessed on 5 March 2026).
48. Central Bank of Brazil. Minutes of the Monetary Policy Committee (Copom) Meeting, July 2025. Available online: https://www.bcb.gov.br/en/pressdetail/2623/nota (accessed on 5 March 2026).
49. Central Bank of Brazil. Minutes of the Monetary Policy Committee (Copom) Meeting, September 2025. Available online: https://www.bcb.gov.br/en/publications/copomminutes/05112025 (accessed on 5 March 2026).
50. Central Bank of Brazil. Minutes of the Monetary Policy Committee (Copom) Central Bank of Brazil. 2022. Available online: https://www.bcb.gov.br/en/publications/copomminutes (accessed on 5 March 2026).