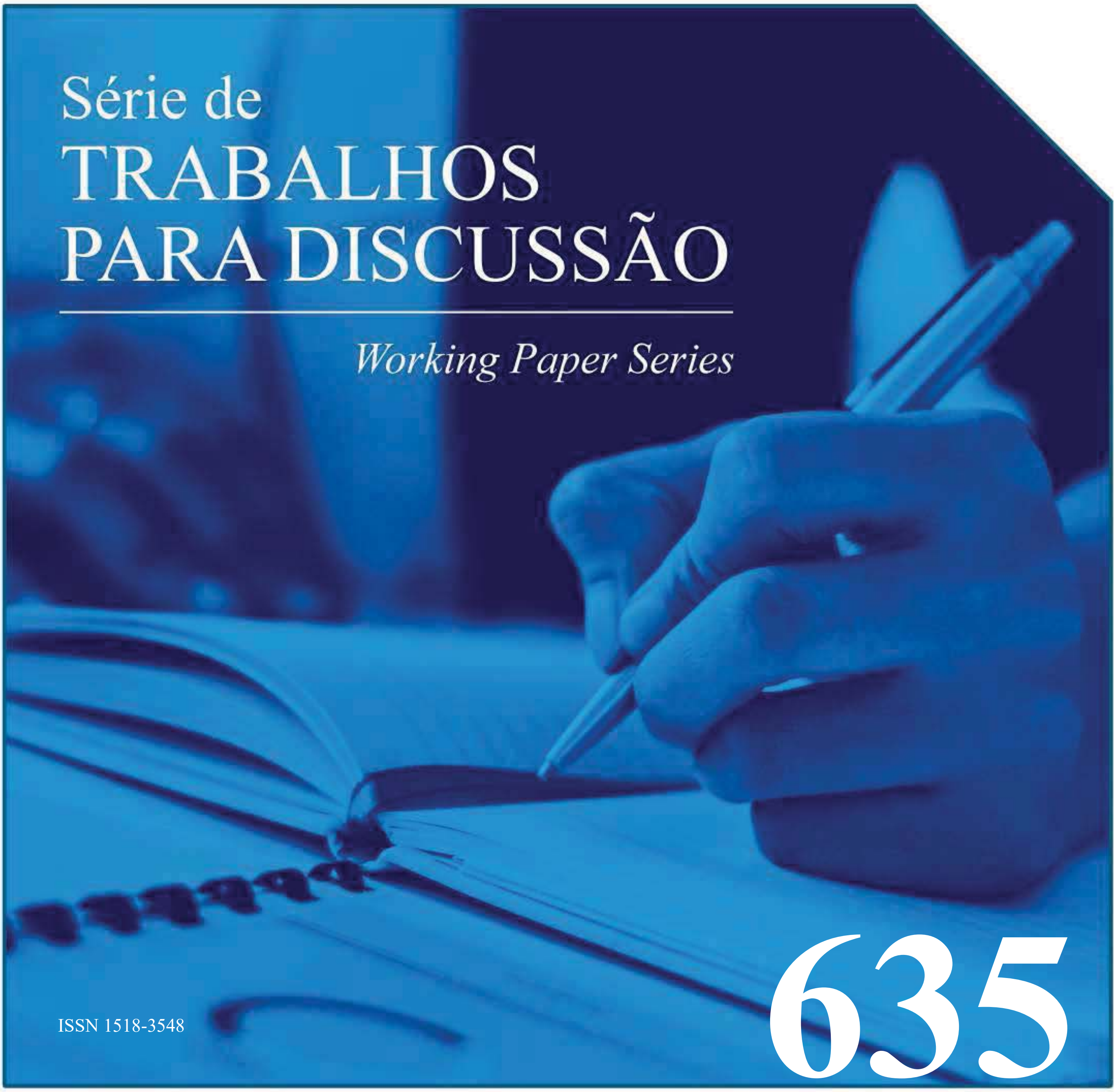

Novembro 2025

The Not So Quiet Revolution: signal and noise in central bank communication

Leonardo Nogueira Ferreira, Caio Garzeri, Diogo Guillen, Antônio Lima, Victor Monteiro

BANCO CENTRAL DO BRASIL

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	no. 635	Novembro	2025	p. 3-29

Working Paper Series

Edited by the Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Rodrigo Barbone Gonzalez

Co-editor: Eurilton Alves Araujo Jr

Head of the Research Department: André Minella

Deputy Governor for Economic Policy: Diogo Abry Guillen

Citizen Service Division

Banco Central do Brasil

Deati/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Toll Free: 0800 9792345

Fax: +55 (61) 3414-2553

Internet: http//www.bcb.gov.br/?CONTACTUS

Non-technical Summary

This paper examines the evolving role of communication in central banking, with a particular focus on the U.S. Federal Reserve. Over the past few decades, central banks have increasingly relied on public statements, meeting minutes, and speeches to guide market expectations and enhance transparency. We investigate whether this growing volume of communication genuinely improves the public's ability to anticipate monetary policy decisions, or whether it sometimes generates more confusion than clarity.

Using advanced text analysis techniques, the study evaluates thousands of Fed communications to determine their predictive value for future interest rate changes. The findings reveal a trade-off between "signal" and "noise." Statements and minutes of meetings by the Federal Open Market Committee (FOMC) and speeches from its Chair and Vice Chair tend to contain meaningful information that helps markets forecast policy moves. In contrast, speeches from other members of the Federal Open Market Committee (FOMC), particularly those without voting power, often lack predictive value and may even obscure the Fed's overall message.

The paper concludes that while transparency remains a vital goal, excessive or poorly coordinated communication can undermine its effectiveness for forecasting purposes. This suggests that central banks might benefit from streamlining their messaging to preserve the clarity and credibility of their guidance.

Sumário Não Técnico

Este artigo examina o papel crescente da comunicação na política monetária, com foco especial banco central dos Estados Unidos. Nas últimas décadas, os bancos centrais passaram a depender cada vez mais de declarações públicas, atas de reuniões e discursos para orientar as expectativas do mercado e promover maior transparência. O artigo investiga se esse aumento no volume de comunicações realmente melhora a capacidade do público de antecipar decisões de política monetária ou se, por vezes, gera mais confusão do que clareza.

Utilizando técnicas avançadas de análise textual, o estudo avalia milhares de comunicações do banco central norte-americano para determinar seu valor preditivo em relação a futuras alterações na taxa de juros. Os resultados revelam um equilíbrio delicado entre "sinal" e "ruído". Os comunicados e atas de reuniões provenientes do comitê (FOMC) e de seu Presidente e do Vice-Presidente tendem a conter informações relevantes que ajudam os mercados a antecipar movimentos da política monetária. Em contraste, os discursos de outros membros do FOMC, particularmente daqueles sem poder de voto, frequentemente carecem de valor preditivo e podem até atrapalhar a mensagem geral do Fed.

O artigo conclui que, embora a transparência continue sendo um objetivo essencial, uma comunicação excessiva ou mal coordenada pode não contribuir para projeções macroeconômicas. Isso sugere que os bancos centrais poderiam se beneficiar de uma abordagem mais enxuta e estratégica em sua comunicação, a fim de preservar a clareza e a credibilidade de suas orientações.

The Not So Quiet Revolution: signal and noise in central bank communication*

Leonardo N. Ferreira[1], Caio Garzeri[2], Diogo Guillen[1], Antônio Lima[1], and Victor Monteiro[3]

[1]Central Bank of Brazil
[2]Genoa Capital
[3]Insper

Abstract

This paper quantifies the "prediction value" of different forms of central bank communication. Combining traditional econometrics and natural language processing, we test how much forecast-improving information can be extracted from the different layers of the Federal Reserve communication. We find that committee-wise communication (statements and minutes) and speeches by the Chair and the Vice Chair improve interest rate forecasts, suggesting that they provide additional information to understand the policy reaction function. However, individual communication beyond the Vice Chair, such as speeches by board members, other FOMC members, and Federal Reserve Bank presidents not sitting in FOMC, is not forecast improving and sometimes even worsens interest-rate forecasts. Based on our theoretical model, we interpret these results as suggesting that the Fed may have overcommunicated, providing excessive noise-inducing communication for forecasting purposes.

JEL Classification: C53, E52, E58

Keywords: Central Bank Communication, Signal, Noise

*We thank Adney Costa Moura for excellent research assistance. We also thank Raphael Schoenle, Yannick Timmer, participants at presentations at the Central Bank of Brazil, the 2024 CEMLA Conference - XXIX Meeting of the Central Bank Researchers Network, the 46th Meeting of the Brazilian Econometric Society, the October 2025 Meeting of the CCA Consultative Group of Monetary Policy of the BIS, and the Workshop on Machine Learning for Economics at Institutions (KOF, BIS, BCB). The views expressed here are those of the authors and do not necessarily reflect those of the Banco Central do Brasil. Corresponding author: leonardo.ferreira@bcb.gov.br.

1 Introduction

Central banks have been increasingly relying on communication to steer the economy. Statements have become less telegraphic and speeches more numerous, addressing current policy issues in a more detailed and frequent manner, making them potentially more relevant to economic agents. Figure 1 provides two dimensions of these trends for the Fed: more depth in a communication event and more numerous communication events.

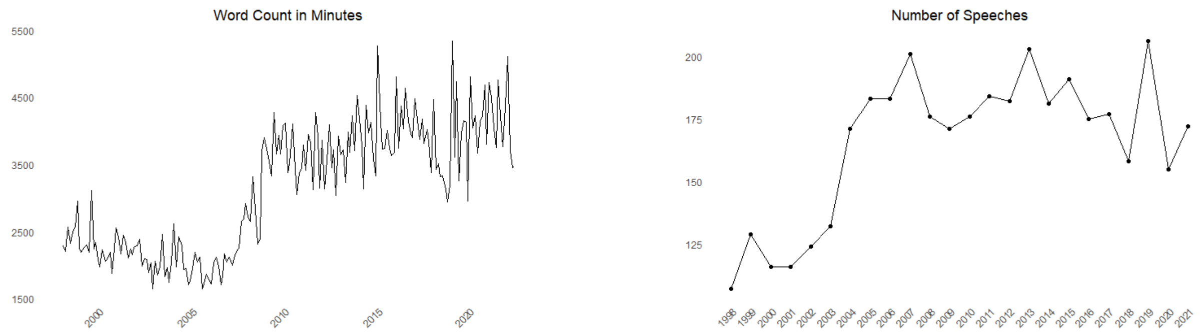

Figure 1: Length of FOMC Minutes and number of Speeches by Fed members over time

Note: Word count excludes stop words. Graph shows the sum of speeches delivered by Fed Chair, Vice Chair, members of the Board of Governors and Reserve Bank presidents.

This revolution in central bank communication poses new challenges both on the policy-making front and backstages (Blinder, 2004, 2018). A critical one is the trade-off between informativeness and distortion. Central bank decision makers' words, as well as their actions, serve as public signals to agents forming expectations (Morris and Shin, 2002).[1] Nonetheless, informing too much can generate unnecessary noise. On the other hand, compressing information too much is also an undesired outcome as it reduces transparency.

In practice, however, the optimal level of communication is yet to be found. On the one hand, Lustenberger and Rossi (2020) conclude that more communication can even increase forecast errors and dispersion. Similarly, Do Hwang et al. (2021) find intensive central bank communication, also measured by the number of speeches, worsens the opinion that executives have of their central bank's impact on the economy. On the other hand, Swanson (2023) and Swanson and Jayawickrema (2023) highlight the importance of Fed Chair speeches as a monetary policy tool and show that, using high-frequency surprises, these speeches are even more important than Federal Open Market Committee (FOMC) announcements for most maturities.

It is worth noting, nonetheless, that these papers do not explore directly the content of the communication. To our knowledge, the only exception is Ahrens and McMahon (2021), who extracted economic signals from central bank speeches. Their initial findings point to

[1] Melosi (2017) adds to this literature by showing the signaling effects of monetary policy can help understand the behavior of inflation and its expectations, but he focuses on interest-rate communication.

the fact that more "cacophonous" communication in the build-up to FOMC meetings might be associated with stronger subsequent market surprises at FOMC policy announcement time.[2]

We, in advancing on this avenue, shed light on the signal-to-noise trade-off, which lies at the heart of the central bank communication. Is it worth communicating? Is it worth communicating divergences among committee members?[3] When does the diversity of views enhance the understanding of the scenario and the policy reaction function, and when does it bring more confusion and misinterpretation, i.e. cacophony (Jefferson, 2024)?[4] Specifically, this paper contributes to this literature by quantifying the "prediction value" of different forms of central bank communication, while taking into account the content of the message and the type of messenger.

We proceed in two steps. First, we test whether the information extracted from different public signals issued by the FOMC and by members of the Federal Reserve System can help predict the path of the fed funds rate. In order to do that, we use Natural Language Processing (NLP) to retrieve information from a variety of documents, and we cumulatively incorporate additional layers of central bank communication – statements, minutes, Chair speeches, Vice Chair speeches, other Board members' speeches, other FOMC members' speeches, and not-sitting-in-FOMC Federal Reserve Bank presidents' speeches – in an otherwise standard Bayesian Direct Forecast (BDF). By doing so, we allow, for instance, the impact of the latter to be different from the impact of speeches delivered by voting members of the FOMC.

Second, we build a simple model of central bank communication, grounded on information theory literature.[5] Unlike signalling models[6] or rational inattention,[7] our framework is based on rate-distortion theory, where central bank optimally transmits information about inflation and the output gap through multiple noisy and potentially correlated communication instruments to minimize the expected distortion in public expectations about the fundamentals, subject to an information constraint à la Shannon capacity. These instruments are not perfect substitutes because their informational content depends on both the precision and the correlation of their noise. In this context, we derive an analytical solution for the optimal central bank communication given by a cutoff rule: the central bank

[2]While not focusing on cacophony, Ahrens et al. (2025) find no evidence that central bankers' speeches resolve uncertainty. In fact, they increase volatility and tail risk in both equity and bond markets.

[3]See Hansen et al. (2014) and Gnan and Rieder (2023) for individual biases and preferences; Blinder (2004), Blinder (2018), Bennani and Neuenkirch (2017) and Tillmann and Walter (2019) for documentation of divergence in monetary policy committees and Vissing-Jorgensen (2019) for policymakers competing for the attention of financial markets.

[4]This cacophony problem has been noted by then-Governor Powell (2016) ("Market participants often say that there are too many voices saying too many different things about policy – the cacophony problem.") and, according to Blinder (2018), will not go away soon. Warsh (2016), for instance, wrote that the Fed "licenses a cacophony of communications in the name of transparency".

[5]Such as Max (1960) and Cover (1999).

[6]One seminal example is Angeletos and Pavan (2007).

[7]See Maćkowiak et al. (2023) for a recent survey.

chooses to remain silent whenever the expected distortion from communication exceeds the benchmark strategy. We also generate testable predictions about the trade-offs involved in using multiple instruments in the central bank's communication strategy.

Results show that the first layers of communication, committee-wise and speeches by the Chair and the Vice Chair, add significant value to forecasts of the fed funds rate in an out-of-sample evaluation exercise. We interpret that statements, minutes and the first layers of speeches improve the understanding of the policy reaction function. However, upon adding further layers that rely on individual members, predictive gains are reversed. Based on our theoretical model, we interpret these results as suggesting that the Fed may have overcommunicated, providing excessive noise-inducing communication for forecasting purposes.

Notably, our findings align with recent survey results on the Fed communication with market participants. According to more than 60% of the Fed watchers surveyed by Wessel and Boocker (2024), speeches by Fed governors and Fed bank presidents are useless or only somewhat useful. On the other hand, almost 90% of them wish the Fed Chair spoke more or the same. The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 presents our empirical framework. Section 4 shows the results. Section 5 introduces a simple model of central bank communication. Section 6 concludes.

2 Data

The benchmark dataset consists of 3 macroeconomic variables from 1998M02, when communication started to become more common, to 2020M02. Core PCE (π) is calculated by taking the first difference of the logarithm of the corresponding index. The unemployment rate (μ) proxies the state of the economy, and the fed funds rate (ffr) is used as a measure of the stance of the Fed. This choice of variables aims to mimic the Fed's reaction function. The series are downloaded from FRED.

2.1 The Corpus of Central Bank Releases

The text-augmented model also includes information retrieved from the FOMC statements that followed scheduled meetings during the period of analysis, the minutes released a few weeks after the policy decision, the speeches delivered by the Fed Chair, the Vice Chair and other members of the Board as well as speeches delivered by the Federal Reserve Bank presidents who were sitting at the FOMC ('in FOMC') and who were not ('not in FOMC') at the time the speeches were given.[8] Statements, minutes and Board members' speeches were retrieved from the Federal Reserve website. Other speeches were scraped from FRASER and regional Fed websites. This gives us a comprehensive dataset of 3,600

[8]Each year, four FOMC votes rotate among 11 Federal Reserve Bank presidents.

speeches, of which 370 were delivered by the Chair and, on the other end, 1,307 were delivered by regional Fed presidents at times they were not filling the rotating seats.[9,10] We turn this information into sentiment scores for the seven layers of communication that are cumulatively incorporated into the different versions of the text-augmented BDF.

2.2 FinBERT and Sentiment Indices

We use FinBERT (Araci, 2019) – Financial BERT –, a BERT-based model (Devlin et al., 2019) trained for financial sentiment analysis tasks.[11] The author shows that this fine-tuning led to a 15% increase in accuracy. During the classification task, FinBERT calculates the probability of three labels for an input text: positive, neutral, and negative. For each document in our corpus, we extract these label probabilities.[12] We consider these probabilities to be the sentiment indices of a given text.

For each set of speeches, we create two monthly time series of sentiment indices: one for the positive score and one for the negative score.[13] In periods with more than one speech in a particular category, we calculate the average of the respective indices. For months with no statements, minutes or speeches in a given layer, we repeat the last value. After this process, we can use these sentiment scores in the regressions as proxies to the sentiment in the different communications of the Federal Reserve System.

3 Empirical Framework

The benchmark model is the BDF proposed by Ferreira et al. (2023). Direct Forecasts (DF) consist of estimating a series of predictive regressions at different horizons of a variable of interest on a set of predictors. BDFs regularize DF regressions via informative priors, producing forecasts that leverage the flexibility of DFs while retaining a degree of estimation uncertainty comparable to Bayesian VARs with standard macroeconomic priors. This approach is ideal for our application in that it does not require iterating communication forward to produce the forecasts of the three macroeconomic variables of interest.[14]

Let y_t denote the n-dimensional vector of endogenous variables at t, and $y_{t+h|t}$ its h-

[9] The Bank for International Settlements (BIS) keeps a database of international central bankers' speeches, but it covers only a subset of the speeches available on the original websites, with just a few in the beginning of our sample or delivered by regional Fed presidents.

[10] The press conference held after each meeting FOMC meeting is also regarded as an important communication tool. Nonetheless, it started only in 2011.

[11] Although the use of LLM may introduce potential "look-ahead" biases (Carriero et al., 2024), this may not be empirically relevant (Araujo et al., 2025). Indeed, an alternative approach that addresses this issue is presented in the appendix and delivers similar results.

[12] More details about this model can be seen in the appendix B.

[13] We use only these scores because the three add up to 1 and the regression has an intercept.

[14] While forecasting communication is an interesting avenue of research, it is beyond the scope of this paper.

step-ahead forecast. The forecasts are computed as follows:

$$y_{T+h|T} = \widehat{\mathbf{B}}^{(h)}\mathbf{y}_T \tag{1}$$

where $\mathbf{y}_T \equiv (1, y'_T, y'_{T-1}, \ldots, y'_{T-p+1})'$, $p = 12$, $h = 3, 6, 12$, and each of the estimated $\widehat{\mathbf{B}}$ matrices of coefficients is of dimension $n \times (np + 1)$. In the benchmark version, $n = 3$. In the seven text-augmented versions, the vector y_t is cumulatively appended with the sentiment scores, ending up with $n = 17$ in the larger model.

To address the increasing number of regressors, we follow Ferreira et al. (2023) and formulate a prior for BDF coefficients that is centered around the coefficients of a VAR with equivalent set of regressors. Hierarchical modeling then allows to optimally select the informativeness of the priors, and the data to optimally deviate from them, at each horizon and for each set of variables.

4 Forecast Evaluation

Models are compared and selected on the basis of their predictive performance. All models are estimated over a rolling data window. Starting from an initial 1998M02-2009M12 window, this results in a set of 111 out-of-sample forecasts. We incorporate layers of communication cumulatively: first with statements, then statements, minutes, and so forth. The comparison is conducted based on root mean squared forecast errors (RMSFE) computed as:

$$RMSFE^i_{t,h} = \sqrt{\frac{1}{P}\sum(\hat{Y}^i_{t+h}(M) - Y^i_{t+h})^2} \tag{2}$$

where $\hat{Y}^i_{t+h}(M)$ denotes the forecast produced by model M for variable i and Y^i_{t+h} is the actual data, and the sum is computed over all the P forecasts produced. Table 1 reports the ratios of the RMSFE relative to the benchmark BDF: values lower than 1 favor the text-augmented BDFs.[15]

Interestingly, the forecasts of the core PCE for $h = 3$ and $h = 6$ do not improve with the inclusion of text. Nonetheless, text seems to be more informative for the forecasts of the core PCE for $h = 12$. The RMSFE ratio is below for all the ratios, although the difference is statistically significant only up to the inclusion of speeches delivered by the Board members. The opposite happens with the unemployment rate forecasts: text adds more value to the forecasts at horizons 3 and 6, but makes forecasts worse for $h = 12$.

Results for the fed funds rate are stronger: text almost always adds value to the predictions, and the improvement is greater compared to the changes in core PCE and the unemployment rate forecasts. Moreover, results exhibit an interesting pattern that is high-

[15]The observations and the forecasts of π are cumulated, so the performance is evaluated based on the inflation over the following quarter, semester and year.

Table 1: Forecast Evaluation: Text-augmented versus benchmark BDF

		Statement	Minute	Chair	V Chair	Board	FOMC	Not FOMC
	3M	1.04	1.03	1.06	1.06	1.07	1.04	1.02
		(0.55)	*(0.51)*	*(0.22)*	*(0.26)*	*(0.29)*	*(0.50)*	*(0.78)*
		(0.27)	*(0.05)*	*(0.00)*	*(0.00)*	*(0.08)*	*(0.03)*	*(0.03)*
π	6M	1.11	1.07	1.05	1.06	1.06	1.04	1.02
		(0.59)	*(0.64)*	*(0.32)*	*(0.24)*	*(0.38)*	*(0.38)*	*(0.29)*
		(0.15)	*(0.04)*	*(0.00)*	*(0.03)*	*(0.01)*	*(0.03)*	*(0.03)*
	12M	0.96	0.94	0.94	0.94	0.93	0.93	0.91
		(0.02)	*(0.41)*	*(0.96)*	*(0.96)*	*(0.93)*	*(0.80)*	*(0.64)*
		(0.00)	*(0.01)*	*(0.13)*	*(0.07)*	*(0.07)*	*(0.11)*	*(0.17)*
	3M	1.03	0.98	0.95	0.93	0.95	0.95	0.93
		(0.49)	*(0.52)*	*(0.55)*	*(0.51)*	*(0.47)*	*(0.64)*	*(0.80)*
		(0.25)	*(0.21)*	*(0.07)*	*(0.01)*	*(0.01)*	*(0.01)*	*(0.03)*
μ	6M	1.11	0.99	0.96	0.94	0.96	0.96	0.95
		(0.11)	*(0.93)*	*(0.62)*	*(0.54)*	*(0.67)*	*(0.71)*	*(0.65)*
		(0.19)	*(0.03)*	*(0.08)*	*(0.09)*	*(0.06)*	*(0.08)*	*(0.07)*
	12M	1.23	1.05	1.03	1.04	1.03	1.03	1.01
		(0.28)	*(0.21)*	*(0.43)*	*(0.20)*	*(0.38)*	*(0.42)*	*(0.42)*
		(0.01)	*(0.01)*	*(0.02)*	*(0.01)*	*(0.02)*	*(0.02)*	*(0.03)*
	3M	0.84	0.92	0.99	0.99	1.01	1.04	1.06
		(0.74)	*(0.55)*	*(0.55)*	*(0.54)*	*(0.55)*	*(0.54)*	*(0.46)*
		(0.06)	*(0.02)*	*(0.02)*	*(0.06)*	*(0.04)*	*(0.05)*	*(0.04)*
ffr	6M	0.86	0.84	0.90	0.83	0.88	0.89	0.90
		(0.14)	*(0.77)*	*(0.88)*	*(0.86)*	*(0.88)*	*(0.89)*	*(0.98)*
		(0.23)	*(0.10)*	*(0.04)*	*(0.02)*	*(0.01)*	*(0.01)*	*(0.01)*
	12M	0.74	0.73	0.74	0.69	0.71	0.76	0.79
		(0.09)	*(0.07)*	*(0.08)*	*(0.03)*	*(0.07)*	*(0.12)*	*(0.18)*
		(0.14)	*(0.16)*	*(0.15)*	*(0.12)*	*(0.18)*	*(0.13)*	*(0.10)*

Notes: The table shows the ratio of the RMSFEs, relative to the benchmark for the 3-, 6- and 12-month-ahead forecasts over 2010M01 2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parentheses. The variables are core PCE (π), the unemployment rate (μ), and the fed funds rate (ffr).

lighted in Figure 2 where we plot the same information from Table 1 but in terms of forecast improvement in order to facilitate visualization.[16] The first layers of communication – statements, minutes and sometimes speeches by the Chair and Vice Chair – show economically and statistically significant improvements in all the horizons.[17] Beyond the Vice Chair, however, forecasts are worse compared to the version of the model estimated only with the statements. In fact, for the 3-month-ahead forecast, the inclusion of speeches

[16]This improvement is computed as $(1 - RMSFE\ ratio) \times 100$.

[17]The positive contribution of the statements to the forecasts of the fed funds rate has already been documented by Ferreira (2021).

delivered by Federal Reserve Bank regional presidents not voting in FOMC worsens forecasts even in comparison with the benchmark BDF.

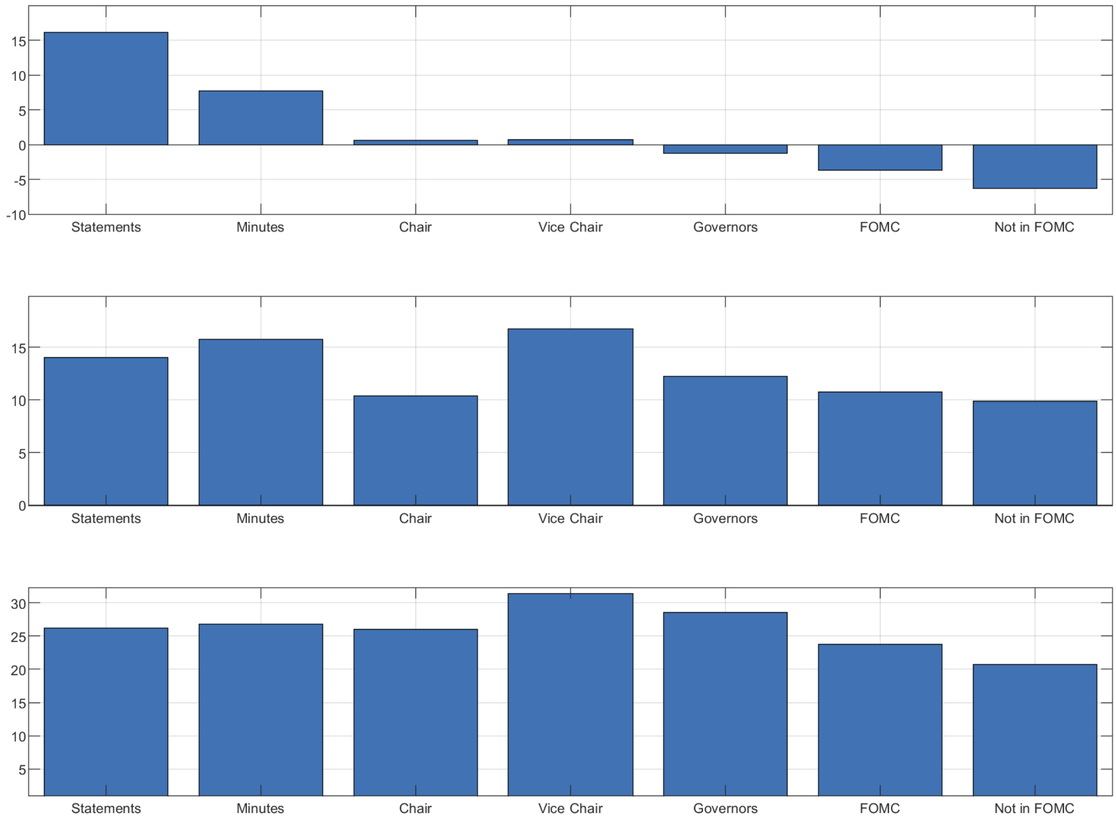

Figure 2: Improvement in the fed funds rate forecast relative to the benchmark for h=3, 6, 12

Overall, the last set of documents adding value to the fed funds rates forecast is the speeches delivered by the Vice Chair, where the optimal level of communication appears to be. Beyond that, there is a consistent decay in predictive performance for all horizons. Interestingly, the inclusion of speeches delivered by the Chair does not seem to improve forecasts beyond the content of statements and minutes. This probably reflects the fact the FOMC is a collegial committee – autocratically collegial during Greenspan's terms and genuinely collegial since Bernanke – and the Chair conveys the position of the consensus, possibly with a personal tweak (Blinder, 2004). The fact that the Chair's main outlet is the press conference held after each FOMC meeting – which is not included in our analysis, as it started only in 2011 – may also be downplaying the importance of the Chair in our results.

Since the predictive power of text – and, consequently, the chosen specification – can change over time, forecast evaluation is also conducted recursively. To this end, we plot a standard loss differential between the larger text-augmented model and the benchmark model, based on squared 12-month-ahead forecast errors. In the same graph, we highlight the periods when Jurado et al. (2015)'s measure of uncertainty exceeds its median.

Interestingly, Figure 3 reveals that the loss differential is higher during times of elevated uncertainty. This suggests that while central bank communication can be seen as particularly valuable during uncertain periods, offering guidance to help economic agents navigate volatile conditions, caution is warranted. It seems that, in uncertain times, central banks could benefit even more from streamlining their message.

Figure 3: Loss differential and uncertainty

Note: The shaded areas identify the periods where Jurado et al. (2015)'s measure of uncertainty exceeds the median. The line depicts a standard loss differential between a benchmark model and the larger text-augmented model, based on their squared 12-month-ahead forecast errors.

Combined, our findings suggest the Fed is probably deviating from the optimal communication strategy. This could happen because the multiplicity of voices in the Federal Reserve System may be creating cacophony and confusing the markets about the "central bank thinking". By controlling for the type of messenger, we shed some light on the apparent contrast in the previous literature. Chair speeches are indeed useful as shown by Swanson (2006), but too much communication can be detrimental as concluded by Lustenberger and Rossi (2020) and Do Hwang et al. (2021). This happens because the increase in the number of speeches usually comes together with an increase in the number of voices.

Finally, given that the improvement in the fed funds rate forecasts is significantly greater than that of core PCE and the unemployment rate forecasts – especially for statements and minutes –, we interpret that the predictive gains for the fed funds rate likely stem from signals about the parameters of policy makers' reaction functions rather than their information advantage. This aligns with the findings in Hoesch et al. (2023), who conclude that information effects are much less important in recent samples.

5 A Simple Model of Central Bank Communication

5.1 An intuitive introduction to the model

Before delving into the model, we first outline the intuition we aim to explore. In our setup, a central bank has a deeper understanding of the economy and decides whether to disclose its information.[18] The question, more so than the content of the message itself, is to evaluate whether it is worth revealing such information. On the one hand, if there is no risk of introducing noise, central banks should always transmit more information. On the other hand, if the risk of getting information lost in translation is too high, they should reveal minimal information. This is the trade-off evaluated here: balancing information revelation with the introduction of noise. One should note, henceforth, that our trade-off is not the usual incentive-provision signalling mechanism; rather, it is more closely related to information theory models, where revealing information inevitably introduces noise.

To illustrate the approach taken here, consider a concrete example. Suppose, for simplicity, that aversion towards inflation has changed, i.e., the coefficient on the Taylor rule changes. Central bankers are not mere machines, and they try to convey such a message saying, for instance, that "inflation is something we should fight aggressively". However, such a statement could be seen as if they are observing an inflationary shock (Phillips curve shock from private information) or as a change on Taylor rule preferences. Obviously, the central bank can always refine its communication, but we focus on the fact that it is just impossible to telegraph messages precisely. Moreover, the more information is revealed, the harder it is to reveal it with the desired precision. The central banker should then weigh the welfare obtained under the prior information on the Taylor coefficient vis-à-vis the welfare implications from providing noisy information. Should it communicate more, but possibly getting lost in translation?

5.2 The Model

Borrowing the environment of the communication mechanism from information theory models, such as Max (1960), we consider an economy where there exist two agents, the central bank and the public. The former is privately informed about the state of the economy and decides to reveal this information under different monetary instruments, while the latter uses this communication to form expectations about the economic environment.

We assume there exist two state variables in this economy that are imperfectly privately observed only by the central bank, inflation, and output gap. They are Gaussian random variables. Subsequently, the central banker decides which instrument – such as statements, minutes, or speeches – should be used to convey information to the public. Each instrument

[18] One should note, however, that this assumption can be easily relaxed by allowing the private signal obtained by the central bank to be uninformative.

has a specific capability to influence agents' expectations, depending on the level of noise in the central bank's communication.

Therefore, there is a trade-off between informativeness and distortion in the communication setting, where central bank incorporates its knowledge about the state variables as well as the expected distortion of its communication strategy. The timeline of the model is given by the following: (i) the central bank observes inflation and output gap; (ii) it chooses the message provided for the set of monetary instruments and; (iii) the public receives central bank information about inflation and output gap.

Economic Environment - We now introduce the primitives that describe the environment and the central bank's decision problem in our model. The central bank observes two latent state variables, inflation π and output gap y, which are jointly normally distributed.

$$(\pi, y) \sim \mathcal{N}(0, \Sigma), \tag{3}$$

where $\mathbb{E}[\pi] = \mathbb{E}[y] = 0$, $\mathbb{V}(\pi) = \sigma_\pi^2$, $\mathbb{V}(y) = \sigma_y^2$, and $\text{Cov}(\pi, y) = \rho\sigma_\pi\sigma_y$.

Consistent with its dual mandate, the central bank cares about communicating its assessment of inflation and output gap to the public. It has access to N communication instruments indexed by $i = \{1, 2, \ldots, N\}$, such as minutes, speeches, statements, to anchor public expectations. Each instrument transmits a noisy signal about both states variables simultaneously.

$$\begin{cases} \hat{\pi}_i = \pi + \varepsilon_i^\pi, \\ \hat{y}_i = y + \varepsilon_i^y, \end{cases} \tag{4}$$

where $(\varepsilon_i^\pi, \varepsilon_i^y)$ are the communication noise terms for instrument i about inflation and output gap, respectively. They are normally distributed with zero mean and variances:$\mathbb{V}(\varepsilon_i^\pi) = \sigma^2_{\varepsilon_i^\pi}$, $\mathbb{V}(\varepsilon_i^y) = \sigma^2_{\varepsilon_i^y}$, $\text{Cov}(\varepsilon_i^\pi, \varepsilon_i^y) = \rho\sigma^2_{\varepsilon_i^\pi}\sigma^2_{\varepsilon_i^y}$. These noise variances are allowed to differ across instruments, and may also be correlated across them, allowing for heterogeneous and overlapping communication precision.

Therefore, this environment allows us to investigate the optimal central bank communication problem under the lens of the monetary policy instruments used to talk to the public. The central banker's desire is to provide the best message to communicate the fundamentals of the economy generating the lowest noise, which here to us is the same as arguing that the central banker minimizes the distortion in his communication every time he talks to the public.

Central Bank Objective Function - The central bank aims to minimize the quadratic expected distortion in the public's understanding of the fundamentals.

$$D = \mathbb{E}\left[w_\pi(\pi - \hat{\pi})^2 + w_y(y - \hat{y})^2\right], \tag{5}$$

where $(\hat{\pi}, \hat{y})$ are the posterior expectations of (π, y) formed by the public using the information conveyed by the instruments, and w_π, w_y denote the relative importance of communication precision in inflation and the output gap, respectively.

To establish the informativeness of central bank communication, we borrow the intuition of mutual information from information theory, incorporating an information-theoretical constraint on the central bank's ability to communicate. The informativeness of the signals is summarized by the mutual information between (π, y) and the set of observed signals $\{\hat{\pi}_i, \hat{y}_i\}_{i=1}^N$:

$$I((\pi, y); \{(\hat{\pi}_i, \hat{y}_i)\}_{i=1}^N) \leq R \tag{6}$$

This constraint captures a communication bottleneck: the central bank cannot fully and perfectly telegraph its private information incorporated in its reaction function to words.

Central Bank Decision Problem - The central bank decision problem is given by the minimization of the deviation of public expectation about the economic fundamentals subject to its limitation to convey information to the public.

$$\begin{aligned} \min_{\{\sigma_\pi^2 \varepsilon_i^\pi, \sigma_y^2 \varepsilon_i^y\}_{i=1}^N} \quad & \mathbb{E}\left[w_\pi(\pi - \hat{\pi})^2 + w_y(y - \hat{y})^2\right] \\ \text{s.t.} \quad & I((\pi, y); \{(\hat{\pi}_i, \hat{y}_i)\}_{i=1}^N) \leq R \end{aligned} \tag{7}$$

Central Bank Communication Strategy - This environment allows us to derive a cutoff rule in which the central bank does not communicate all the time. Proposition 1 describes that when the distortion arising from noisy communication exceeds the benchmark communication strategy, the central bank optimally chooses not to communicate.

Proposition 1 *Assuming that the central bank uses N instruments to communicate to the public and the communication optimality condition is given by the inequality $D \leq D^*(R)$, where the distortion function, given by equation 5, could be redefined as follows:*

$$D = w_\pi \cdot \left(\frac{1}{\sigma_\pi^2} + \mathbf{1}_N^\top Cov(\boldsymbol{\varepsilon}^\pi)^{-1}\mathbf{1}_N\right)^{-1} + w_y \cdot \left(\frac{1}{\sigma_y^2} + \mathbf{1}_N^\top Cov(\boldsymbol{\varepsilon}^y)^{-1}\mathbf{1}_N\right)^{-1} \tag{8}$$

where $\mathbf{1}_N^\top$ is the indicator function that represents the instruments used by the central bank, D the effective distortion function and $D^*(R)$ is the optimal distortion provided by central bank's benchmark strategy.

Then, the optimal communication strategy is defined by the following cutoff rule:

$$R(D) = \begin{cases} \log\left(\frac{2\sqrt{w_\pi w_y \sigma_\pi^2 \sigma_y^2}}{D}\right), & if \quad D \leq D^* \\ 0, \quad if \quad D > D^* \end{cases} \tag{9}$$

The Proposition 1 defines a rate-distortion rule: the central bank communicates only if its effective distortion lies below or equal the benchmark level D^*. The effective rate distortion function $R(D)$ is a function of three main terms. The noise structure of the instruments, the weights of the reaction functions, and the ability of the public to extract information from central bank communication.

This cutoff rule highlights the informational role of instruments in the transmission channel between the central bank's reaction function and the expectations of the public. Naturally, it raises the same question that we investigated in the empirical counterpart: When is it optimal to communicate through a single instrument rather than multiple ones? Proposition 2 sheds light that in periods of high uncertainty in communication, it is better to use few instruments.

Proposition 2 *Given the economic environment from equations 3, 4, 5, and 6. For instruments sufficiently correlated there exists an equilibrium in which the central bank strictly prefers to communicate through a single instrument rather than through a combination of multiple instruments.*

The Proposition 2 emphasizes that when communication conveyed by different instruments is highly correlated, the marginal informational value of adding layers of instruments can be negative or null. In this case, the optimal strategy is to use with the highest precision.

In our model, the use of more instruments may reduce the distortion if the signals are independent and convey new information. However, if the noises are positively correlated, the combination of instruments delivers redundant information, increasing the distortion in the communication, and potentially worsening the public inference about central bank reaction function.

In line with our empirical results, where the addition of layers in the central bank communication could enhance the noise sent in the message, we propose a simulation of our theoretical model. We simulate an economy where the central bank has access to three instruments to communicate the fundamentals: statements, minutes, and chair speeches. We sequentially compute the expected distortion as the instruments are added. For simplicity, we assume that the third instrument is highly correlated with the previous ones. [19]

[19] We assume a simple calibration where correlation between the instruments is equal to one. The noise issued by the statement is 0.1, for minutes we assume 0.5 and for chair speeches it is equal to 200000.

Figure 4 describes the resulting distortion path. It shows that while adding minutes reduces the distortion, the inclusion of chair speeches fails to generate further improvement. The message from chair speeches is nearly redundant and adds little or no informational content, consistent with our theoretical prediction. Therefore, in this example, the communication layer of the chair speeches should not be used, corroborating our empirical results.

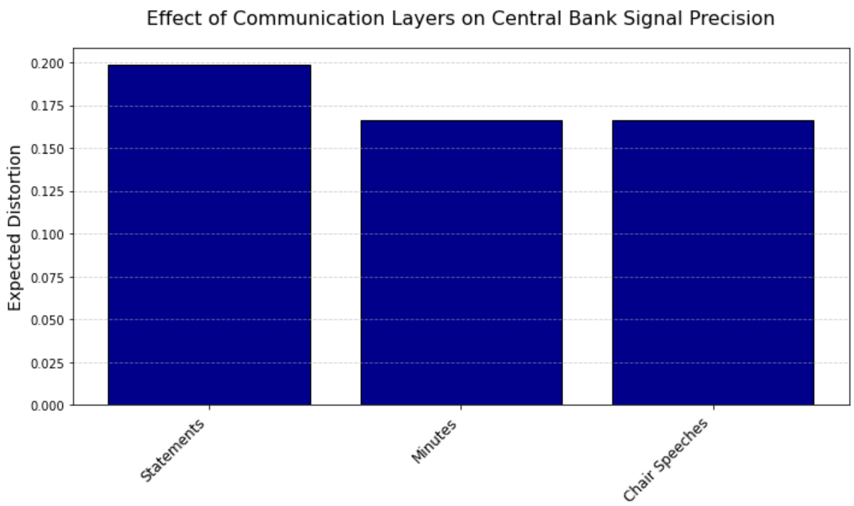

Figure 4: Optimal Central Bank Ratio-Distortion Function

6 Conclusion

We have empirically tested how much forecast-improving information can be extracted from the public signals issued by the Federal Reserve, while taking into account the content of the message and the type of messenger. Results indicate that committee-based communication and speeches by the Chair and the Vice Chair add significant value to forecasts of the fed funds rate in an out-of-sample evaluation exercise. However, individual communication beyond the Vice Chair reverses these predictive gains. Based on our theoretical model, we interpret these findings as suggesting that the Fed may have overcommunicated, providing excessive noise-inducing communication for forecasting purposes. It is worth highlighting, however, that central bank communication has multiple objectives, and providing guidance to economic agents forming their expectations is just one of them. In this vein, speeches by regional Fed presidents may not add value to forecasting, in line with the surveyed by Wessel and Boocker (2024), but may serve different purposes such as transparency and accountability.

References

Ahrens, M., Erdemlioglu, D., McMahon, M., Neely, C. J., and Yang, X. (2025). Mind your language: Market responses to central bank speeches. *Journal of Econometrics*, 249:105921.

Ahrens, M. and McMahon, M. (2021). Extracting economic signals from central bank speeches. In *Proceedings of the Third Workshop on Economics and Natural Language Processing*. Association for Computational Linguistics.

Angeletos, G.-M. and Pavan, A. (2007). Efficient use of information and social value of information. *Econometrica*, 75(4):1103–1142.

Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.

Araujo, D., Bokan, N., Comazzi, F., and Lenza, M. (2025). Word2prices: embedding central bank communications for inflation prediction.

Bennani, H. and Neuenkirch, M. (2017). The (home) bias of European central bankers: new evidence based on speeches. *Applied Economics*, 49(11):1114–1131.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan):993–1022.

Blinder, A. S. (2004). *The quiet revolution: Central banking goes modern*. Yale University Press.

Blinder, A. S. (2018). Through a crystal ball darkly: The future of monetary policy communication. In *AEA Papers and Proceedings*, volume 108, pages 567–571. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Carriero, A., Pettenuzzo, D., and Shekhar, S. (2024). Macroeconomic forecasting with large language models. *arXiv preprint arXiv:2407.00890*.

Cover, T. M. (1999). *Elements of information theory*. John Wiley & Sons.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). 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, volume 1 (long and short papers)*, pages 4171–4186.

Do Hwang, I., Lustenberger, T., and Rossi, E. (2021). Does communication influence executives' opinion of central bank policy?. *Journal of International Money and Finance*, 115:102393.

Ferreira, L. N. (2021). *Forecasting with VAR-teXt and DFM-teXt Models: exploring the predictive power of central bank communication.* Banco Central do Brasil.

Ferreira, L. N., Miranda-Agrippino, S., and Ricco, G. (2023). Bayesian local projections. *The Review of Economics and Statistics*, pages 1–45.

Gado, K. (2024). Transformers, central banker speeches, tail stock returns.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6):1545–1578.

Gnan, P. and Rieder, K. (2023). The (not so) quiet period: Communication by ECB decision-makers during monetary policy blackout days. *Journal of International Money and Finance*, 130:102744.

Hansen, S. and McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99:S114–S133.

Hansen, S., McMahon, M., and Rivera, C. V. (2014). Preferences or private assessments on a monetary policy committee? *Journal of Monetary Economics*, 67:16–32.

Hoesch, L., Rossi, B., and Sekhposyan, T. (2023). Has the information channel of monetary policy disappeared? revisiting the empirical evidence. *American Economic Journal: Macroeconomics*, 15(3):355–387.

Jefferson, P. N. (2024). Communicating about monetary policy: A speech at central bank communications: Theory and practice," a conference hosted by the Federal Reserve Bank of Cleveland, Cleveland, Ohio., May 13, 2024. Technical report.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.

Lustenberger, T. and Rossi, E. (2020). Does central bank transparency and communication affect financial and macroeconomic forecasts? *62nd issue (March 2020) of the International Journal of Central Banking.*

Maćkowiak, B., Matějka, F., and Wiederholt, M. (2023). Rational inattention: A review. *Journal of Economic Literature*, 61(1):226–273.

Max, J. (1960). Quantizing for minimum distortion. *IRE Transactions on Information Theory*, 6(1):7–12.

Melosi, L. (2017). Signalling effects of monetary policy. *The Review of Economic Studies*, 84(2):853–884.

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

Powell, J. (2016). A view from the Fed. *Speech at the" Understanding Fedspeak" event cosponsored by the Hutchins Center on Fiscal and Monetary Policy at the Brookings Institution and the Center for Financial Economics at Johns Hopkins University, Washington, DC*.

Swanson, E. and Jayawickrema, V. (2023). Speeches by the fed chair are more important than fomc announcements: An improved high-frequency measure of us monetary policy shocks. *Unpublished Manuscript*.

Swanson, E. T. (2006). Have increases in Federal Reserve transparency improved private sector interest rate forecasts? *Journal of Money, Credit and Banking*, pages 791–819.

Swanson, E. T. (2023). The importance of Fed chair speeches as a monetary policy tool. *AEA Papers and Proceedings*, 113:394–400.

Tillmann, P. and Walter, A. (2019). The effect of diverging communication: The case of the ECB and the Bundesbank. *Economics Letters*, 176:68–74.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

Vissing-Jorgensen, A. (2019). Central banking with many voices: The communications arms race.

Warsh, K. (2016). The Federal Reserve needs new thinking. *Wall Street Journal*, 8(24):2016.

Wessel, D. and Boocker, S. (2024). Federal Reserve communications: Survey results. *Hutchins Center on Fiscal and Monetary Policy at Brookings*.

Appendix

A. Theoretical Results

In this appendix, we provide the proofs of the propositions of the Section of The Model. Starting with the proof of Proposition 1.

Proposition 1:

The result of Proposition 1 is the cutoff rule of the central bank decision problem. To prove it, we start characterizing the economic environment, deriving the optimal rate-distortion function of the central bank D^*, which characterizes the lower feasible distortion in the communication of the central bank. Then, we derive the cutoff rule.

In order to derive the optimal rate-distortion function, we assume that the state variables, inflation π and the output gap y, are independently Gaussian even though the signals are correlated, and the central bank minimizes a quadratic and separable distortion function under a mutual information constraint, such as: $\pi \sim \mathcal{N}(0, \sigma_\pi^2)$, $y \sim \mathcal{N}(0, \sigma_y^2)$, $\text{Cov}(\pi, y) = \rho\sigma_\pi\sigma_y$. These assumptions guarantee the existence of a lower bound for the rate-distortion function. Then, by equation 5, the central bank's distortion function can be stated as the following:

$$D = w_\pi \mathbb{E}[(\pi - \hat{\pi})^2] + w_y \mathbb{E}[(y - \hat{y})^2] = w_\pi D_\pi + w_y D_y \tag{10}$$

Let R denote the mutual information between the true state (π, y) and the signal $(\hat{\pi}, \hat{y})$.

Under Gaussianity and separability assumption, we can decompose the mutual information according to each state variable, as given by the following:

$$R = R_\pi + R_y, \tag{11}$$

Where R_π and R_y are the information used to encode π and y, respectively.

As derived by Cover (1999), due to the Gaussian distribution to the state variables and separability assumption of the rate-distortion function, we can write the terms as the following:

$$\begin{aligned} D_\pi &= \sigma_\pi^2 e^{-2R_\pi} \\ D_y &= \sigma_y^2 e^{-2R_y} \end{aligned} \tag{12}$$

Therefore, substituting equation 12 into 10, the central bank minimization problem can be written as the following:

$$\begin{aligned} &\min_{R_\pi, R_y} w_\pi \sigma_\pi^2 e^{-2R_\pi} + w_y \sigma_y^2 e^{-2R_y} \\ &\text{subject to } R_\pi + R_y = R. \end{aligned} \tag{13}$$

Lagrangian can be written as follows:

$$\mathcal{L} = w_\pi \sigma_\pi^2 e^{-2R_\pi} + w_y \sigma_y^2 e^{-2R_y} + \lambda(R_\pi + R_y - R) \tag{14}$$

Taking the first-order condition with respect to R_π and R_y, we get the following:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial R_\pi} &= -2w_\pi \sigma_\pi^2 e^{-2R_\pi} + \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial R_y} &= -2w_y \sigma_y^2 e^{-2R_y} + \lambda = 0 \end{aligned} \tag{15}$$

Equating the two expressions for λ:

$$\begin{aligned} &2w_\pi \sigma_\pi^2 e^{-2R_\pi} = 2w_y \sigma_y^2 e^{-2R_y} \implies \\ &e^{-2R_\pi} = \frac{w_y \sigma_y^2}{w_\pi \sigma_\pi^2} e^{-2R_y} \end{aligned} \tag{16}$$

Let's define: $A = w_\pi \sigma_\pi^2$ and $B = w_y \sigma_y^2$. Now, taking the logarithm of equation 16.

$$R_\pi = \frac{1}{2} \log\left(\frac{A}{B}\right) + R_y \tag{17}$$

Substituting equation 17 into the constraint of the central bank minimization problem in the equation 10, we get the optimal level of noise, such as the following:

$$R = \frac{1}{2} \log\left(\frac{A}{B}\right) + 2R_y \tag{18}$$

Now, when we use the constraint that $R = R_\pi + R_y$, and solve for R_y, we obtain the following equation:

$$R_\pi + \left(R_\pi - \frac{1}{2} \log \frac{A}{B}\right) = R \tag{19}$$

Solving for R_π:

$$R_\pi = \frac{R}{2} + \frac{1}{4} \log \frac{A}{B} \tag{20}$$

Analogously, for R_y:

$$R_y = \frac{R}{2} - \frac{1}{4}\log\frac{A}{B} \tag{21}$$

Now, plugging into distortion equation 10, we obtain the following:

$$D(R) = w_\pi \sigma_\pi^2 e^{-R}\left(\frac{A}{B}\right)^{\frac{1}{2}} + w_y \sigma_y^2 \left(\frac{A}{B}\right)^{\frac{1}{2}} \tag{22}$$

Then, substituting the terms A and B, rearranging the terms of the equation 22, the optimal distortion-rate function can be written:

$$D(R) = 2\sqrt{w_\pi w_y \sigma_\pi^2 \sigma_y^2} \cdot e^{-R} \tag{23}$$

And the inverse rate-distortion function is:

$$R(D^*) = \log\left(\frac{2\sqrt{w_\pi w_y \sigma_\pi^2 \sigma_y^2}}{D^*}\right) \tag{24}$$

Therefore, the central bank only convey information whenever the effective distortion D is less noisy than the optimal D^*. It means that central bank communication strategy can be written by the following cut rule:

$$R(D) = \begin{cases} \log\left(\frac{2\sqrt{w_\pi w_y \sigma_\pi^2 \sigma_y^2}}{D}\right), & \text{if} \quad D \leq D^* \\ 0, \quad \text{if} \quad D > D^* \end{cases} \tag{25}$$

Now, we provide the proof for Proposition 2.

Proposition 2:

Assuming the economic environment described in the equations 3, 4, 5, and 6, we can prove this proposition by induction. First, when we consider the central bank uses a unique instrument to communicate, the distortion associated with this instrument i, with noise variances $\sigma^2_{\varepsilon_i^\pi}$ and $\sigma^2_{\varepsilon_i^y}$.

We consider that instrument i conveys information from both state variables, such that inflation and output gap transmitted by the central bank is defined by the following equations:

$$\begin{aligned} \hat{\pi}_i &= \pi + \varepsilon_i^\pi \\ \hat{y}_i &= y + \varepsilon_i^y. \end{aligned} \tag{26}$$

Without loss of generality, let's assume that variance of the errors terms are invariant in the time as well as the correlation between them according to the use of communication instruments. Then, for any pair of instruments i and j, we get:

$$\mathbb{V}(\varepsilon_i^\pi) = \mathbb{V}(\varepsilon_i^y) = \sigma^2, \quad \mathrm{Cov}(\varepsilon_1^\pi, \varepsilon_2^\pi) = \mathrm{Cov}(\varepsilon_1^y, \varepsilon_2^y) = \rho, \quad \mathrm{Cov}(\varepsilon_i^\pi, \varepsilon_j^y) = 0 \text{ for all } i, j.$$

In this context, the noise covariance matrix can be written as the inner product between an indicator function I and the correlation between the state variables of the economy, such as in the following matrix:

$$\Sigma_\varepsilon = I_2 \otimes \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \tag{27}$$

Let $W = \mathrm{diag}(w_\pi, w_y)$ be the weight matrix for distortion, and suppose without loss of generality $w_\pi = w_y = 1$. The result of this proposition holds for all positive combinations of weights between the interval $[0, 1]$ as the distortion function is a convex combination.

Thus, the posterior variance of the case when the central bank uses a single instrument can be written as follows:

$$D_1 = \left(\frac{1}{\sigma_\pi^2} + \frac{1}{\sigma^2}\right)^{-1} + \left(\frac{1}{\sigma_y^2} + \frac{1}{\sigma^2}\right)^{-1} \tag{28}$$

Now, to derive the posterior variance of the case when the central bank uses two instruments to communicate to the public, we consider using two instruments, i and j, with symmetric noise σ^2 and correlation ρ in the signal errors.

$$\mathrm{Cov}(\boldsymbol{\varepsilon}^\pi) = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \quad \Rightarrow \quad (\mathrm{Cov}(\boldsymbol{\varepsilon}^\pi))^{-1} = \frac{1}{(1-\rho^2)\sigma^2} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix} \tag{29}$$

By the separability property of the rate-distortion function, we can break the analysis in two terms, one to the inflation and other to the output gap. The scalar precision of the joint signal according to the inflation can be written as:

$$\mathbf{1}_2^\top (\mathrm{Cov}(\boldsymbol{\varepsilon}^\pi))^{-1} \mathbf{1}_2 = \frac{1 - 2\rho + 1}{(1-\rho^2)\sigma^2} = \frac{2(1-\rho)}{(1-\rho^2)\sigma^2} \tag{30}$$

So the posterior variance becomes:

$$\mathrm{Var}(\pi \mid \hat{\pi}_1, \hat{\pi}_2) = \left(\frac{1}{\sigma_\pi^2} + \frac{2(1-\rho)}{(1-\rho^2)\sigma^2}\right)^{-1} \tag{31}$$

As $\rho \to 1$ the term $\frac{2(1-\rho)}{(1-\rho^2)} \to 1$. Therefore, the posterior variance when the central bank uses two instruments converge to:

$$\mathrm{Var}(\pi \mid \hat{\pi}_1, \hat{\pi}_2) \to \left(\frac{1}{\sigma_\pi^2} + \frac{1}{\sigma^2}\right)^{-1} = \mathrm{Var}(\pi \mid \hat{\pi}_1) \tag{32}$$

The derivation of the posterior variance of the output gap follows the same intuition.

When we combine both cases, we have that rate distortion function when the central bank uses two instruments converge to the same distortion function when the central bank uses a single instrument whenever ρ converge to one, i.e., $D_2 \to D_1 \quad \text{as } \rho \to 1$.

Moreover, if the second instrument has strictly larger noise or is more correlated, then $D_2 \geq D_1$. Thus, there exists a threshold correlation $\bar{\rho} \in (0, 1)$ such that for all $\rho \geq \bar{\rho}$ $D_2 \geq D_1$.

By induction whether the central bank uses uses N to convey information, such that $N > 2$, whenever the addition of a new instrument k, with noise highly correlated ρ with the existing $k-1$ instruments, such that $\rho > \bar{\rho}$, the distortion strictly increases. Therefore, in the limit, the rate distortion function under k instruments must be strictly higher than under $k-1$ instruments, such that $D_k > D_{k-1}$.

In the last part of the proof, we must show that this also holds for $N = k+1$ instruments.

Adding $k+1$ instrument with correlated noise $\rho \geq \bar{\rho}$, we can apply the result from Sherman–Morrison–Woodbury identity, which guarantees that the effective increase in precision from the inverse covariance after the inclusion of a highly collinear signal is negligible. It holds because the total precision of the central bank communication when it uses k instruments to communicate P_k is tantamount to the sum of the covariances between the instrument due to the separability of the rate distortion-function, as described by:

$$P_k = \sum_{i=1}^{k} Cov(\epsilon_i)^{-1} \tag{33}$$

Then, when central bank adds a new instrument $k+1$ in its communication strategy, the precision becomes:

$$P_{k+1} = P_k + Cov(\epsilon_{k+1})^{-1} \tag{34}$$

In the case where $Cov(\epsilon_{k+1})^{-1}$ is nearly collinear to the previous instruments, it means that the contribution of this instrument in the communication of the central bank could be approximated by a rank-1 matrix. Therefore, by Sherman-Morrison-Woodbury identity, whenever the signal is redundant, the updated inverse precision matrix changes only marginally.

In addition, using the fact that the marginal gain in Fisher information from a highly correlated signal diminishes as correlation increases, where in the extreme case, the marginal Fisher information approaches to zero. We have that the following condition holds:

$$\mathbf{I}(k+1)^\top \mathrm{Cov}(\boldsymbol{\varepsilon}^\pi)^{-1} \mathbf{I}(k+1) \approx \mathbf{I}k^\top \mathrm{Cov}(\boldsymbol{\varepsilon}^\pi)^{-1} \mathbf{I}k \tag{35}$$

That is, adding a highly correlated instrument adds almost no new precision and may worsen conditioning. Thus:

$$D_{k+1} \geq D_k \tag{36}$$

Therefore, by induction, the result holds for all N such that $\rho \geq \bar{\rho}$.

In order to characterize this proof, we provide two examples that parametrize the environment of our economy and provide the economic intuition behind Proposition 2.

B. Transformer, BERT and FinBERT models

The transformer model as described by Vaswani et al. (2017) has an encoder-decoder structure. The encoder maps an input sequence of symbol representations $(x_1, ..., x_n)$ to a sequence of continuous representations $z = (z_1, ..., z_n)$. Given z, the decoder then generates an output sequence $(y_1, ..., y_n)$ of symbols one element at a time. At each step, the model is autoregressive, consuming the previously generated symbols as additional input when generating the next.

The Bidirectional Encoder Representations from Transformers (BERT) model architecture is a multilayer bidirectional transformer encoder based on the original transformer implementation proposed by Vaswani et al. (2017). BERT uses a masked language model (MLM) pretraining objective, inspired by the Cloze task (Devlin et al., 2019). The MLM objective enables the representation to fuse the left and the right contexts, which allows the possibility to pre-train a deep bidirectional transformer. In addition to the MLM, a next sentence prediction task was jointly used to pre-train a text-pair representation.

BERT was pre-trained in a corpus of books (800M words) and in the English Wikipedia Corpus (2500M words). Once the model has been pre-trained, it can be used on specific NLP tasks. Apart from output layers, the same architecture is used in both pre-training and fine-tuning steps. The same pre-trained model parameters are used to initialize models for different downstream tasks. During the fine-tuning step, all parameters are fine-tuned for the specific task. The result of this approach was that BERT advanced the state-of-the-art for eleven NLP tasks.

The FinBERT model, in turn, is a BERT-based language model trained for financial NLP tasks (Araci, 2019). The author implemented further pre-training the BERT model on a financial domain corpus. The corpus used in this further pre-training was the TRC2-financial. It is a subset of Reuters' TRC24, which consists of 1.8M news articles that were published by Reuters between 2008 and 2010. The main sentiment analysis dataset used for fine-tuning was Financial PhraseBank, which consists of 4,845 English sentences selected randomly from financial news found in the LexisNexis database. Another dataset used for sentiment analysis was the FiQA Sentiment that was created for the WWW '18 conference financial opinion mining and question-answering challenge.

In both data sets used for financial sentiment analysis, FinBERT achieved state-of-the-art results by a significant margin. For example, in the classification task, the model increased accuracy by 15%. These results provide empirical evidence that the FinBERT model is good enough for extracting explicit sentiments in the representation of scores such as the sentiment indices we used in this paper. As in Gado (2024), we use the first n words of a document FinBERT can take without batching.

C. Additional Results

This section presents the results based on a VAR model augmented with topic proportions calculated for the same corpus. We apply Latent Dirichlet Allocation (LDA) to the meeting minutes in order to compute the topic-specific term probabilities (Blei et al., 2003). In order to make the analysis more precise, this first step is conducted at the level of the paragraph. While this approach overlooks the sentiment of the text, it addresses look-ahead bias since topic-specific term probabilities are computed using minutes only up to 2012.

Then, keeping the topic-specific term probabilities fixed at their estimated values for minutes' paragraphs, we estimate aggregate document distributions for minutes, statements and different types of speeches. By doing that, we can focus on the part of the speeches that are related to monetary policy. Such topic proportions are the time series incorporated in the VAR.

Table 1 below presents the 10 most common terms per estimated topic over the minutes. As in Hansen and McMahon (2016), K will be set to 15. We select 6 of them - Topics 6, 8, 10,11, 12, 14 - which are more closely related to discussions about the economic situation. After analyzing these words, we consider these topics to be covering mainly the following themes: **Topic 6**: Housing Market; **Topic 8**: Output; **Topic 9**: Inflation; **Topic 11**: Risk; **Topic 12**: Monetary Policy; **Topic 15**: Labor Market.

Table 1: 10 most common terms per topic

Topic 6	**Topic 8**	**Topic 9**	**Topic 11**	**Topic 12**	**Topic 15**
level	spend	price	particip	committe	month
hous	consum	inflat	econom	polici	increas
remain	busi	expect	note	monetari	employ
sale	invest	energi	risk	condit	averag
low	incom	increas	outlook	stabil	rate
home	household	core	meet	feder_fund_rate	unemploy
rate	recent	cost	term	percent	rose
mortgag	report	consum	longer	maintain	gain
continu	expenditur	measur	financi	sustain	end
activ	confid	recent	general	consist	labor

These topic time series are then cumulatively incorporated into the BVAR. The shape of the plot displaying the results is similar. For the first two layers of communication – statements and minutes – the text-augmented VARs show significant improvements in the forecasts of the fed funds rate for all horizons. However, speeches by other members of the FOMC and presidents of Reserve Banks not sitting in the FOMC consistently worsen forecasts.

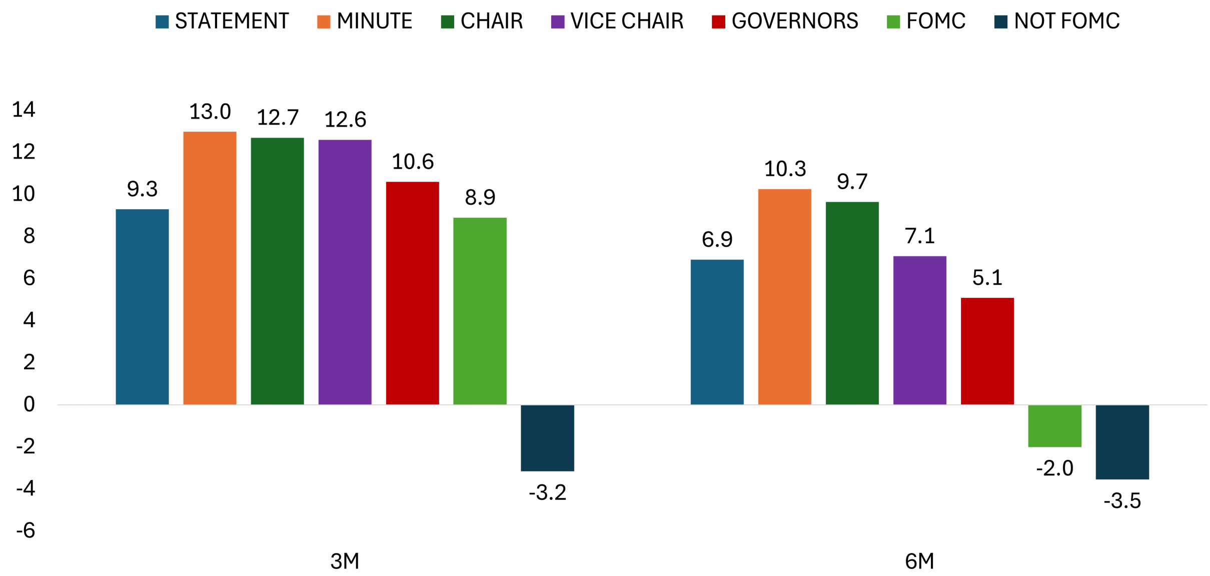

Figure 1: Improvement in the fed funds rate forecast in relation to the benchmark

Note: The bars show the difference in predictive log-scores multiplied by 100, relative to the benchmark for the 3-month and 6-month-ahead forecasts. Different colors represent models with increasing layers of communication.