The not so quiet revolution: Signal and noise in Fed communication*

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

[1]Central Bank of Brazil
[2]PUC Rio
[3]Ministry of Planning and Budget

This version: July 19, 2024.

Abstract

This paper quantifies the "prediction value" of different forms of central bank communication. Combining traditional econometrics and textual machine learning techniques, 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) improves interest-rate forecasts up to six months ahead, suggesting it provides additional information to understand the policy reaction function. However, individual communication such as speeches by the Vice Chair, board members, other FOMC members, and Federal Reserve Bank presidents not seating in FOMC are not forecast improving and sometimes even worsen the interest-rate forecast. Based on our theoretical model that policy makers trade-off providing more information but inherently adding noise when communicating, we interpret these results as suggesting that the Fed may have over-communicated, providing excessive noise-inducing communication.

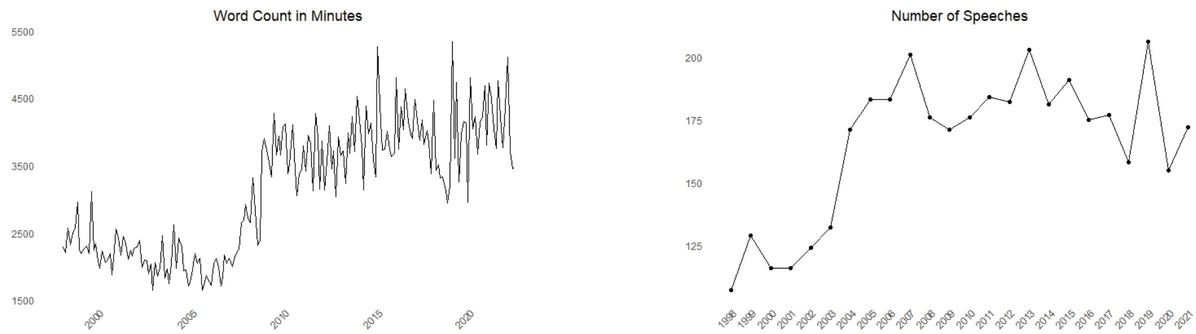
JEL Classification: C53, E52, E58
Keywords: Central Bank Communication, Signal, Noise, Cacophony

*We thank Adney Costa Moura for excellent research assistance. The views expressed here are those of the authors and do not necessarily reflect those of the Banco Central do Brasil or Ministerio do Planejamento. 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, being potentially more relevant for economic agents. Figure 1 provides two dimensions of such trends for the Fed: more depth on 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 described in section 3. 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 to the policymaking 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 quantity of speeches, worsens the opinion executives have of their central bank's impact on the economy. On the other hand, Swanson (2023) highlights the importance of Fed Chair speeches as a monetary policy tool and shows, 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 have not explored directly the content of the communication.

We, on exploring it, shed light on the signal-to-noise trade-off dimension which lies at the heart of the central bank communication. Is it worth communicating? Is it worth

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

communicating divergences among committee members?[2] 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)?[3]

This paper contributes to this literature by quantifying the "prediction value" of different forms of central bank communication, while taking into account the type of messenger and the content of the message. Additionally, we contribute by building a comprehensive dataset on speeches delivered by members of the Federal Reserve System since 1998, including Federal Reserve Bank presidents, which can be used for future research.

We proceed in two steps. First, we present a simple model of central bank communication that helps understand the theoretical optimal communication strategy and the trade-off inherent to central bank communication. Then, we test whether the information extracted from the different public signals issued by the FOMC and by members of the Federal Reserve System are consistent with this strategy and can help predict the path for the fed funds rate. Specifically, 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-seating-in-FOMC Federal Reserve Bank presidents' speeches – in an otherwise standard Bayesian VAR.

Results show that the first layers of communication, committee-wise, add significant value to forecasts of the fed funds rate in an out-of-sample evaluation exercise. We interpret that statements and minutes improve the understanding of the policy reaction function. However, upon adding further layers that rely on individual members, predictive gains are reversed. We interpret these results as over-transparency, inducing noise enhancing communicaiton.

Interestingly, our results corroborate recent results obtained from a survey 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 introduces a simple model of central bank communication. Section 3 describes the data. Section 4 briefly describes our empirical framework. Section 5 presents the results. Section 6 concludes.

[2]See Hansen et al., 2014; Gnan and Rieder, 2023 for individual biases and preferences; Blinder, 2004, 2018; Bennani and Neuenkirch, 2017; 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

[3]This cacophony problem has been noted by then-Governor Powell (2016)[4] 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".

2 A Simple Model of Central Bank Communication

2.1 An intuitive introduction to the model

Before we delve into the formal theory of the model proposed, we start with the intuition that we want to explore herein.

A central bank in our setup has a better understanding of the economy and decides whether to reveal its information [5]. The question, more so than the content of the message itself of the information revealed, is to evaluate whether it is worth revealing such information. On the one hand, if there was no risk of introducing noise, it should always transmit. On the other hand, if the risk of getting information lost in translation is too high, it should reveal little information. This is the trade-off evaluated here: adding noise but revealing information. One should note, henceforth, that our trade-off is not the usual incentive-provision signalling mechanism, it is much more related to information theory models, in which it is impossible to reveal information without introducing noise.

Let us give a concrete example to stress how the approach here should be understood. Suppose, for simplicity, that the preference of aversion towards inflation has changed, i.e., the coefficient on the Taylor rule changes. Central bankers are not pure machines, and they try to convey such message saying, for instance, that "inflation is something we should fight aggressively". However, this 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 shape the communication to try to reveal information in the most precise way. We will abstract from that, assume it is done in such a way, and focus on the fact that it is just impossible to reveal it precisely. Moreover, the more information is revealed, the harder it is to reveal it precisely. The central banker should then compare the welfare obtained under the prior information on the Taylor coefficient vis-à-vis the welfare from providing noisy information. Should it communicate, but possibly getting lost in translation?

2.2 The Model

Borrowing the environment of 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 agent is privately informed about the state of the economy and decides to reveal his information under different monetary instruments, while the latter uses this communication to update his belief and form expectations about the economic environment.

We assume that inflation is a Gaussian random variable, such that $\pi \sim \mathcal{N}(0, \sigma)$. The

[5] One should note, though, that this assumption is easily relaxed by allowing the private signal obtained by the central bank to be uninformative

central bank observes inflation, and, after that, decides about the instrument, such as statements, minutes and speeches, that should be used to reveal information to the public. Each instrument has a specific capability to affect the expectations of the agents, depending on the level of noise in central bank communication.

Therefore, there is a trade-off between informativeness and distortion in the communication setting. Central bank decisions incorporate its knowledge about inflation 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 imperfectly inflation; (ii) it chooses the message provided for the set of monetary instruments and; (iii) the public receives central bank information about inflation.

Central Bank Environment - Now, we set the primitives of our model, describing the environment of central bank decision problem. We consider that central bank can use Z different instruments to communicate the conduct of monetary policy. For each one of them, there exists a distortion function $d_i(\pi, \hat{\pi})$ drawn as a squared-error function as the following:

$$d_i(\pi, \hat{\pi}) = (\pi - \hat{\pi})^2 \tag{1}$$

where $\hat{\pi}$ is the inflation imperfectly communicated by the central bank, given the instrument used.

This equation describes that the distortion in the communication of the central bank is a function of how much revealed inflation deviates from the actual inflation. Intuitively, this equation provides a flavor on the informativeness of central bank communication for a given instrument. Naturally, the central bank desires to fully-anchor the expectations of the agents, which means that its communication with the public should be as precise as possible.

We define the average distortion $D(\pi, \hat{\pi})$ as the expected value of the combination of distortion functions for all instruments used by the central bank. Moreover, we also consider that there is a maximum average level of distortion accepted by the central bank given by D^*.

$$D(\pi, \hat{\pi}) = E[\sum_{i=i}^{Z} \alpha_i d_i(\pi, \hat{\pi})] \tag{2}$$

where $\sum_{i=1}^{Z} \alpha_i = 1$, such that α_i gives the weight of the instrument i in the average distortion of communication.

To set the informativeness of central bank communication, we borrow the intuition from mutual information concept from information theory. Therefore, the average informativeness $I(\pi, \hat{\pi})$ quantifies the amount of information captured by the public about inflation given the information revealed by the central bank.

$$I(\pi, \hat{\pi}) = H(\hat{\pi}) - H(\hat{\pi}|\pi) = H(\pi) - H(\pi|\hat{\pi}) \tag{3}$$

where $H(\cdot)$ captures the entropy. The term $H(\hat{\pi})$ measures the prior uncertainty of the signal about inflation and $H(\hat{\pi}|\pi)$ measures the remaining uncertainty after the communication of the central bank, which is the noise.

Central Bank Decision Problem - Considering the environment of the central bank decision, the optimal communication strategy is obtained through the following minimization problem. The central bank maximizes the informativeness of its communication setting such that the level of noise provided does not exceed the maximum level of distortion accepted.

$$\begin{aligned} &R(D^*) = \min_{p(\hat{\pi}|\pi)} I(\pi, \hat{\pi}) \\ &s.t.\ D(\hat{\pi}, \pi) \leq D^* \end{aligned} \tag{4}$$

Central Bank Communication Strategy - The optimal decision of the central bank about its communication strategy is given by the ratio-distortion function $R(D)$. This equation describes the informativeness of the messages provided in its communication setting given the level of noise provided. Therefore, optimal strategy is given as following:

$$R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{D^*} \text{ if } 0 \leq D^* \leq \sigma^2 \\ 0 \text{ if } D^* > \sigma^2 \end{cases} \tag{5}$$

The intuition here is the following: whenever the noise provided in the communication is lower than the informativeness of the central bank, the optimal decision is to communicate with the public. Otherwise, the optimal policy is doing nothing. Therefore, this equation opens the possibility that multiple instruments delivering messages about monetary policy, with different levels of informativeness, could generate excessive noise in the communication of the central bank, increasing the level of D^*, and cause too much distortion for the level of information provided.

In the following sections, we dive into the empirical counterpart of the paper to test whether the intuition provided in the theoretical model holds for US data. Empirically, we explore different instruments and voices in Fed communication to test whether the Fed is deviating from the optimal communication strategy, talking too little or excessively and sending too much noise.

3 Data

The benchmark dataset consists of 3 macroeconomic and 1 financial variables from 1998M09, when the release of statements after scheduled meetings arose, to 2020M02. Industrial production growth and consumer price inflation are calculated by taking the first difference of the logarithm of the corresponding indices downloaded from FRED-MD (McCracken and Ng, 2016). The fed funds rate is used as a measure of the stance of the Fed. Finally, to capture information about future economic activity and improve the informational content of the regression, the excess bond premium (EBP), introduced by Gilchrist and Zakrajšek (2012), is also incorporated into the vector of variables. The EBP is a corporate bond credit spread purged from the default risk, a useful leading indicator. This choice of variables aims to mimic a central bank reaction function adjusted for variations in spreads in line with Curdia and Woodford (2010).

3.1 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 seating at the FOMC ('in FOMC') and who were not ('not in FOMC') at the time the speeches were given.[6] Statements, minutes and Board members' speeches were retrieved from the Federal Reserve website. Other speeches were scrapped from FRASER and regional Fed websites. This gives us a comprehensive dataset of 3,600 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.[7,8]

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. 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.[9] By doing that, we can focus on the part of the speeches that are related to monetary policy.[10] Such topic proportions are the time series which will be incorporated

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

[7]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.

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

[9]We use minutes only up to 2012 to compute the topic-specific term probabilities, therefore avoiding the use future information that would compromise the forecasting exercise.

[10]In Appendix D we alternatively apply LDA to speeches in order to compute the topic-specific term probabilities and then use them in the other documents. Results are consistent with our main findings.

in the VAR.[11]

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. Those are marked in blue. By analyzing these words we consider these topics to be discussing 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.[12]

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

4 Empirical Framework

The benchmark VAR model follows:

$$Y_t = \sum_{j=1}^{p} \beta_j Y_{t-j} + \mu + A_0 \varepsilon_t \tag{6}$$

where Y_t is the $N \times 1$ vector of standard macro variables, p denotes the lags, with $p = 1, ..., P$, and A_0 is a decomposition of the covariance matrix Σ such that $Var(u_t) = A_0 A_0' = \Sigma$. In the text-augmented VAR, the vector Y_t is appended with the topic time series.

5 Forecast Evaluation

Models are compared and selected based on their predictive performance. All the models are estimated recursively over an expanding data window. Starting from an initial

[11]The topic proportions are transformed into first differences and standardized as in Larsen and Thorsrud (2019). For months with more than one document of the same type, we take averages of topic proportions among documents. For months with no document of a given type, topic proportions are repeated from the previous month with an existing document.

[12]More details of the corpus and of the LDA estimation are presented in Appendix B.

1998M09-2012M12 window, this results in a set of 80 out-of-sample forecasts. We incorporate layers of communication cumulatively: first with statements, then statements and minutes and so forth. The comparison is conducted based on log scores. Given the draws for β and Σ , it is straightforward to compute the 1-, 3-, and 6-step ahead forecast densities by simulating Y_t forward:

$$H(\hat{Y}_{t+h}/Y_t) = \int H(\hat{Y}_{t+h}/Y_t, \Gamma) \times H(\Gamma/Y_t) d\Gamma \tag{7}$$

where $h = 1, 2, 3, ...$ and $\Gamma = \{\beta, \Sigma\}$.

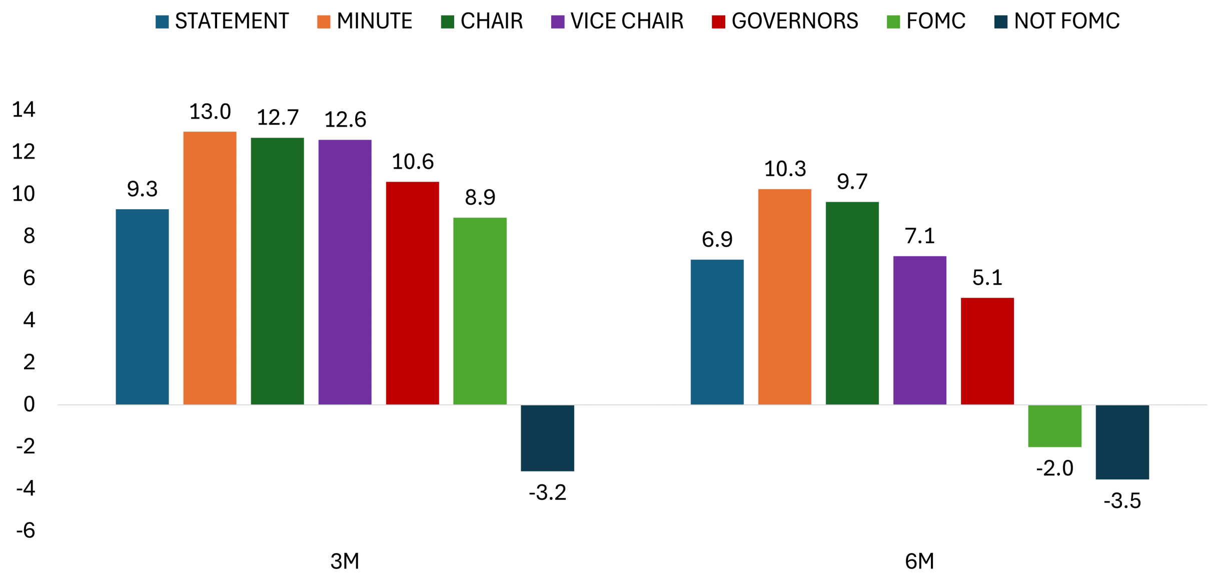
Log-scores are then the (log) likelihood the model assigns to the actual observations Y_{t+1} given data up to t:

$$LS^i_{t,h} = \ln H(Y^i_{t+h}/Y_t) \tag{8}$$

where both Y^i_{t+h} and Y_t are actual data. Following Alessandri and Mumtaz (2017), these (log) predictive densities are estimated using kernel methods.

Figure 2 reports the average differences in log-scores over the entire evaluation period between the augmented VAR with increasing layers of communication and the benchmark VAR: positive values favor the text-augmented models. Table 2 shows that most differences are significant.

Figure 2: Average Log-Score Difference 3-month and 6-month ahead forecasts



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.

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.[13] We interpret such results as a positive contribution of the committee-wise written communication to the understanding of the policy reaction function.

Table 2: Forecast Evaluation: Text-augmented versus benchmark VAR

	Statement	Minute	Chair	Vice Chair	Governors	FOMC	Not FOMC
3M	9.32	12.99	12.73	12.59	10.61	8.89	-3.15
	(0.008)	*(0.031)*	*(0.182)*	*(0.651)*	*(0.660)*	*(0.625)*	*(0.248)*
	(0.000)	*(0.000)*	*(0.000)*	*(0.003)*	*(0.036)*	*(0.008)*	*(0.002)*
6M	6.89	10.26	9.66	7.07	5.10	-2.20	-3.55
	(0.360)	*(0.602)*	*(0.567)*	*(0.197)*	*(0.108)*	*(0.069)*	*(0.005)*
	(0.000)	*(0.000)*	*(0.001)*	*(0.049)*	*(0.026)*	*(0.044)*	*(0.108)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 3-month and 6-month ahead forecasts over 2013M01–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis.[14]

The inclusion of speeches delivered by the Chair does not change much the results, probably reflecting 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). We test this overlapping by comparing the forecasts of a VAR which includes only Chair speeches (without previous layers) with the forecasts produced by the benchmark VAR. Table 3 shows the words delivered by the Chair do add value to forecasts, even though their contribution is lower than when joined by the messages telegraphed in statements and minutes.

Table 3: Forecast Evaluation: Text-augmented versus benchmark VAR - Chair Speeches without any other layer

	3M	6M
Chair	6.70	6.75
	(0.008)	*(0.218)*
	(0.000)	*(0.000)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 3-month and 6-month ahead forecasts over 2013M01–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis.

However, forecasts get worse when adding other layers of communication: speeches by the Vice Chair and by the governors impair predictability for all horizons. Speeches by

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

[14] Such p-values, however, are only indicative since the distributions of the tests are derived based on fixed rolling window estimators.

other members of the FOMC and presidents of Reserve Banks not seating in the FOMC consistently worsen forecasts. Combined, such results 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 by 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.

In Appendix D, we provide additional results and robustness checks across many dimensions. First, they do not depend on the Chair in charge. Second, results for the post zero lower bound period are similar. Third, we find that statements and minutes improve marginally the forecast of inflation and industrial production and only for some horizons. This is consistent with the findings in Hoesch et al. (2023), which conclude that information effects are much less important in recent samples. Therefore, the predictive gains for the fed funds rate seem to come from signals about the parameters of policy makers' reaction functions rather than their information advantage.

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 show that committee-based communication add significant value to forecasts of the fed funds rate in an out-of-sample evaluation exercise. However, individual communication reverts the predictive gains. We interpret these results as suggesting that the Fed may have over-communicated, providing excessive noise-inducing communication.

References

Alessandri, P. and Mumtaz, H. (2017). Financial conditions and density forecasts for US output and inflation. *Review of Economic Dynamics*, 24:66–78.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92.

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.

Curdia, V. and Woodford, M. (2010). Credit spreads and monetary policy. *Journal of Money, credit and Banking*, 42:3–35.

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.

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

Gilchrist, S. and Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4):1692–1720.

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.

Griffiths, T. L. and Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(suppl 1):5228–5235.

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 Prat, A. (2018). Transparency and deliberation within the FOMC: a computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870.

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.

Larsen, V. H. and Thorsrud, L. A. (2019). The value of news for economic developments. *Journal of Econometrics*, 210(1):203–218.

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*.

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

McCracken, M. W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

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.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

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. 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.

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. Full Model

In order to solve the optimization of central bank decision, we start restating the problem here as the following:

$$
\begin{aligned}
& R(D) = \min_{p(\hat{\pi}|\pi)} I(\pi, \hat{\pi}) \\
& s.t.\ D(\hat{\pi}, \pi) \leq D^*
\end{aligned}
\tag{9}
$$

The constraint about the distortion provided in the central bank communication can be rewritten, such that the problem becomes:

$$
\begin{aligned}
& R(D) = \min_{p(\hat{\pi}|\pi)} I(\pi, \hat{\pi}) \\
& s.t.\ \sum_{i}^{Z} \alpha_i \sum_{\hat{\pi},\pi} p(\hat{\pi}, \pi) d_i(\hat{\pi}, \pi) \leq D^*
\end{aligned}
\tag{10}
$$

Now, we calculate the lower bound solution for the decision problem. Recall that $H(\pi) = -\int_{\pi} f(\pi) \log f(\pi) d\pi$.

$$
\begin{aligned}
& I(\pi, \hat{\pi}) = H(\pi) - H(\pi|\hat{\pi}) \\
& I(\pi, \hat{\pi}) = \frac{1}{2}\log(2\pi\pi e)\sigma^2 - H(\pi - \hat{\pi}|\hat{\pi}) \geq \frac{1}{2}\log(2\pi\pi e)\sigma^2 - H(\pi - \hat{\pi}) \\
& I(\pi, \hat{\pi}) \geq \frac{1}{2}\log(2\pi\pi e)\sigma^2 - H(\mathcal{N}(0, E[(\pi - \hat{\pi})^2])) \\
& I(\pi, \hat{\pi}) \geq \frac{1}{2}\log(2\pi\pi e)\sigma^2 - \frac{1}{2}\log(2\pi e)E[(\pi - \hat{\pi})^2]) \\
& I(\pi, \hat{\pi}) \geq \frac{1}{2}\log(2\pi\pi e)\sigma^2 - \frac{1}{2}\log(2\pi e)D^* \\
& I(\pi, \hat{\pi}) \geq \frac{1}{2}\log(\frac{\sigma^2}{D^*})
\end{aligned}
\tag{11}
$$

Therefore, the lower bound is given by:

$$
I(\pi, \hat{\pi}) = \frac{1}{2}\log(\frac{\sigma^2}{D^*}) \tag{12}
$$

Implicitly to the derivation of the lower bound, we are considering that maximum entropy D^* for fixed second order moments is fixed. Now, for achievability, if we have $D^* > \sigma^2$, we choose $\hat{\pi} = 0$, and achieve $R(D^*) = 0$. Otherwise, if we consider another random variable Y, such that $Y \sim \mathcal{N}(0, D^*)$, and we have that $\pi = \hat{\pi} + Y$, for $\hat{\pi}$ and Y

independent, the mutual information of this channel matches the lower bound. Therefore, the rate distortion function decreases with the level of distortion, until it get one, which means no informativeness on central bank communication.

B. Corpus

LDA is estimated with a collapsed Gibbs sampling algorithm. First, the estimation of word's topic assignment consists of the following steps:[15]

Step 1. Randomly allocate to each token in the corpus a topic assignment drawn uniformly from $\{1, ..., K\}$, where K denotes the number of topics.

Step 2. For each token, sequentially draw a new topic assignment via multinomial sampling where

$$P[q_{d,n} = k | Q^{-}, W, \alpha, \eta] \propto \frac{m^{-}_{k,v} + \eta}{\sum_v m^{-}_{k,v} + V\eta}(n^{-}_{d,k} + \alpha) \tag{13}$$

where the '-' superscript denotes counts excluding (d, n) term, with d representing documents and n their terms. $q_{d,n} = k$ denotes the topic assignment of (d, n) term and $Q = (\mathbf{q}_1, ..., \mathbf{q}_D)$ the other topic assignments. $W = (\mathbf{w}_1, ..., \mathbf{w}_D)$ is the observed data. α is the prior on the document-specific mixing probabilities and η on the topic-specific term probabilities. $m_{k,v} \equiv \sum_n \sum_d 1(q_{d,n} = k)1(w_{d,n} = v)$ is the number of times topic k allocation variables generate term v, and $n_{d,k} \equiv \sum_n 1(q_{d,n} = k)$ is the number of words in document d that have topic allocation k. V is the number of unique terms.

Step 3. Repeat Steps 1 and 2 until the required number of draws has been reached.[16]

To fit the LDA, the number of topics K and hyperparameters α and η need to be fixed a priori. Following Griffiths and Steyvers (2004) and Hansen et al. (2018), η is set to $200/V$ and α to $50/K$. Additionally, terms which with sparsity of at least 99.95% are removed (i.e., terms not found in 99.95% of the documents in the Corpus).

Moreover, to avoid using future information in the forecasting, we estimate the topics using the sample only up to December 2012. Then, keeping the topic-specific term probabilities φ fixed at their estimated values, we estimate aggregate document distributions for all documents (statements, minutes, and speeches) for the entire sample. This is done

[15] This follows very closely the Online Appendix of Hansen et al. (2018). For more details, see also Murphy (2012).

[16] This was implemented using the 'topicmodels' package in R (available on https://cran.r-project.org/web/packages/topicmodels/index.html).

by sequentially sampling from:

$$P[\tilde{q}_{d,n} = k | \tilde{Q}^{-}, \tilde{W}, \alpha, \eta] \propto \hat{\varphi}_{v_{d,n}} (n_{\tilde{d},k}^{-} + \alpha) \tag{14}$$

where tilde denotes the new document level. Nonetheless, this gives estimates of each word's topic assignment, since the topic proportions θ were integrated out in the derivation of the collapsed Gibbs sampling. In order to recover them, the output of interest, the final step is to compute the document predictive distributions using:

$$\tilde{\hat{\theta}}_{\tilde{d}}^{k} = \frac{n_{\tilde{d},k} + \alpha}{\sum_{k=1}^{K} n_{\tilde{d},k} + \alpha} \tag{15}$$

Figure 4 below presents the 10 most common terms per estimated topic over the minutes. As detailed in Section 3 Hansen and McMahon (2016), K was set to 15 and 6 topics were selected because they most closely related to discussions about the economic situation.

Table 4: 10 most common terms per topic

Topic 1	**Topic 2**	**Topic 3**	**Topic 4**	**Topic 5**	**Topic 6**	**Topic 7**	**Topic 8**
econom	effect	growth	period	market	level	member	spend
import	demand	expect	intermeet	secur	hous	econom	consum
foreign	substanti	rate	market	purchas	remain	current	busi
export	like	econom	yield	system	sale	risk	invest
trade	mani	project	chang	account	low	might	incom
dollar	firm	forecast	treasuri	agenc	home	outlook	household
declin	well	anticip	index	committe	rate	polici	recent
valu	also	meet	littl	open	mortgag	prospect	report
major	increas	staff	term	direct	continu	like	expenditur
good	reduc	next	equiti	develop	activ	develop	confid

Topic 9	**Topic 10**	**Topic 11**	**Topic 12**	**Topic 13**	**Topic 14**	**Topic 15**
price	growth	particip	committe	inventori	bank	month
inflat	continu	econom	polici	product	market	increas
expect	remain	note	monetari	equip	credit	employ
energi	pace	risk	condit	manufactur	financi	averag
increas	moder	outlook	stabil	vehicl	loan	rate
core	indic	meet	feder_fund_rate	motor	liquid	unemploy
cost	recent	term	percent	industri	commerci	rose
consum	suggest	longer	maintain	busi	condit	gain
measur	slow	financi	sustain	sector	fund	end
recent	econom	general	consist	real	debt	labor

C. VAR Estimation

Following Bańbura et al. (2010), equation (6) can be written in a more compact way, which is more convenient for the estimation:

$$Y = Z\beta + \varepsilon A_0 \tag{16}$$

where $Y = (Y_1, ..., Y_T)'$, $Z = (Z_1, ..., Z_T)'$ with $Z_t = (Y'_{t-1}, ..., Y'_{t-P}, X'_{t-1}, 1)'$, $\varepsilon = (\varepsilon_1, ..., \varepsilon_T)'$, and $\beta = (\beta_1, ..., \beta_P, \phi, \mu)'$ is the $(NP + K) \times N$ matrix containing all the coefficients.

A conjugate prior for the VAR parameters is then introduced by augmenting the vector of variables with the following dummy observations:

$$Y_{d,1} = \begin{pmatrix} \frac{diag(\delta_1\sigma_1,...,\delta_N\sigma_N)}{\tau} \\ 0_{N(P-1)\times N} \\ \cdots\cdots\cdots \\ diag(\sigma_1, ..., \sigma_N) \\ \cdots\cdots\cdots \\ 0_{1\times N} \end{pmatrix}, \quad Z_{d,1} = \begin{pmatrix} \frac{J_p \otimes diag(\sigma_1,...,\sigma_N)}{\tau} & 0_{NP\times 1} \\ \cdots\cdots\cdots & \\ 0_{N\times NP} & 0_{N\times 1} \\ \cdots\cdots\cdots & \\ 0_{1\times NP} & c \end{pmatrix} \tag{17}$$

where δ_1 to δ_N denote the prior mean for the coefficients on the first lag, τ controls the overall tightness of the prior distribution on the VAR coefficients, c is the tightness of the prior on the intercept and the exogenous variables and $J_p = diag(1, ..., P)$ to denote the lags, with $P = 13$. The prior means are chosen as the OLS estimates of the coefficients of an AR(1) regression estimated for each endogenous variable, and σ_i's are set using the standard deviation of the error terms from these regressions. As is standard for US data, $\tau = 0.1$ and $c = 1/10^5$ indicating an informative prior on the lags of the endogenous variables but flat for the intercept and the exogenous variables.

As highlighted by Bańbura et al. (2010), the literature suggests the forecasting performance can be improved by adding a prior on the sum of the coefficients of the form:

$$Y_{d,2} = \left(\frac{diag(\delta_1\mu_1,...,\delta_N\mu_N)}{\lambda} \right), \quad Z_{d,2} = \left(\frac{1_{1\times P} \otimes diag(\sigma_1,...,\sigma_N)}{\lambda} \quad 0_{N\times 1} \right) \tag{18}$$

where μ_i denotes the sample means of the endogenous variables. Following Bańbura et al. (2010), the tightness of the prior of the sum of coefficients is set to $\lambda = 10\tau$, a loose prior.

These vectors are plugged into the vector with the actual observations. Given the artificial data, the observables and the product of the LDA, the model is estimated using an MCMC sampler. The algorithm cycles through the following steps:

Step 1. Draw Σ from the Inverse Wishart:

$$H(\Sigma/Y, X) = IW(S^*, T^*)$$

where the posterior scale matrix is given by $S^* = (Y^* - Z^*\beta^*)'(Y^* - Z^*\beta^*)$, with $\beta^* = (Z^{*'}Z^*)^{-1}(Z^{*'}Y^*)$, $Y^* = [Y; Y_{d,1}; Y_{d,2}]$ and $Z^* = [Z; Z_{d,1}; Z_{d,2}]$. T^* denotes the posterior degrees of freedom given by the number of rows of Y^*.

Step 2. Draw β from the Normal distribution:

$$H(\beta/\Sigma, Y, X) = N(\beta^*, \Sigma \otimes (Z^{*'}Z^*)^{-1})$$

Step 3. Repeat steps 1 to 2 until the required number of draws has been reached.

In this application, the algorithm is iterated 7,000 times for each data window, with the first 4,000 draws discarded as burn-in. This gives 3,000 draws, which are used to simulate the posterior distributions of β and Σ.

D. Additional Results

Our results hold when restricting the sample to the post ZLB period. In Table 5 forecasts for the period from December 2015 to February 2020 are show. Figure 3 also shows that results do not depend on the Fed Chair.

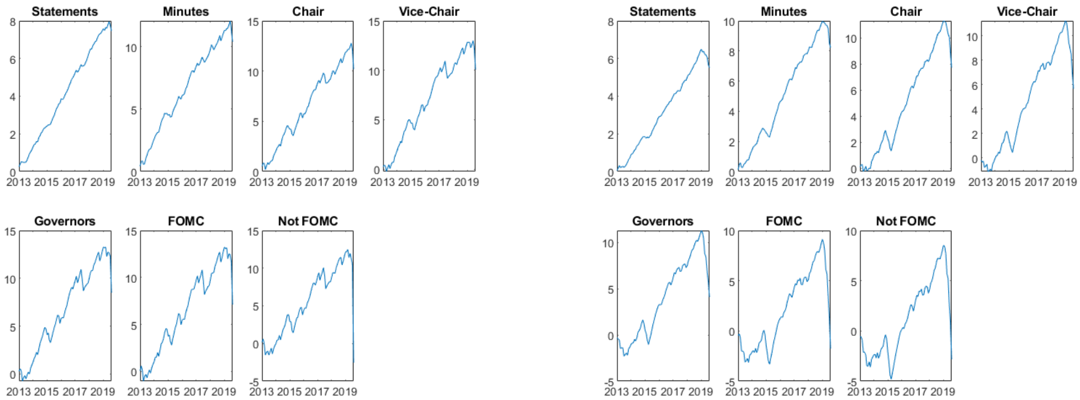
Full results for our main exercise are shown in Table 6 below. Forecasts for the fed funds rate (i), inflation rate (π), output growth(y) and EBP(s) generally follow the same pattern. The addition of statements will improve forecasts; minutes may also improve forecasts depending on the variable and forecast horizon. Speeches by the FED Chair, Vice-Chair, Governors and others members of the FOMC and chairs of the Reserve Banks however will worsen the forecasts, although with less statistical significance.

Table 5: Forecast Evaluation: Text-augmented versus benchmark VAR - Post ZLB period

	Statement	Minute	Chair	Vice Chair	Governors	FOMC	Not FOMC
3M	8.58	9.99	9.74	7.95	5.36	2.51	-16.28
	(0.000)	*(0.005)*	*(0.018)*	*(0.085)*	*(0.331)*	*(0.195)*	*(0.453)*
	(0.000)	*(0.000)*	*(0.000)*	*(0.008)*	*(0.015)*	*(0.007)*	*(0.024)*
6M	6.44	8.88	8.63	4.60	3.10	-5.49	-5.22
	(0.000)	*(0.000)*	*(0.002)*	*(0.007)*	*(0.020)*	*(0.042)*	*(0.060)*
	(0.000)	*(0.000)*	*(0.005)*	*(0.011)*	*(0.017)*	*(0.039)*	*(0.032)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 3-month and 6-month ahead forecasts over 2015M12–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis.

Figure 3: Cumulative Log-Score Difference over Time - 3-month (left) and 6-month (right) ahead forecasts



Note: The lines show the cumulative difference in log-scores between the VAR-text and the VAR models with increasing layers of communication for horizons 3 and 6.

In addition to that, we conducted an alternative exercise for robustness in which we estimate the LDA over speeches and then use these probabilities over all other documents. Results are similar to our main exercise and are shown in Table 7.

Table 6: Forecast Evaluation: Text-augmented versus benchmark VAR

	i			π			y			s		
	1M	3M	6M	1M	3M	6M	1M	3M	6M	1M	3M	6M
Statement	8.18	9.32	6.89	2.14	0.91	-6.79	2.71	3.43	2.08	4.89	4.63	7.53
	(0.000)	*(0.008)*	*(0.360)*	*(0.031)*	*(0.073)*	*(0.795)*	*(0.212)*	*(0.419)*	*(0.000)*	*(0.440)*	*(0.000)*	*(0.000)*
	(0.000)	*(0.000)*	*(0.000)*	*(0.005)*	*(0.010)*	*(0.002)*	*(0.050)*	*(0.011)*	*(0.000)*	*(0.127)*	*(0.000)*	*(0.000)*
Minute	11.78	12.99	10.26	3.02	1.98	-14.67	-1.83	0.85	2.11	-0.62	7.23	7.82
	(0.000)	*(0.031)*	*(0.602)*	*(0.093)*	*(0.841)*	*(0.705)*	*(0.684)*	*(0.509)*	*(0.001)*	*(0.239)*	*(0.000)*	*(0.007)*
	(0.007)	*(0.000)*	*(0.000)*	*(0.746)*	*(0.116)*	*(0.291)*	*(0.507)*	*(0.373)*	*(0.071)*	*(0.142)*	*(0)*	*(0)*
Chair	4.59	12.73	9.66	-20.53	-40.11	-50.72	-8.03	-0.99	-3.91	-5.38	6.22	9.32
	(0.004)	*(0.182)*	*(0.567)*	*(0.584)*	*(0.873)*	*(0.354)*	*(0.227)*	*(0.225)*	*(0.611)*	*(0.264)*	*(0.008)*	*(0.113)*
	(0.000)	*(0.000)*	*(0.001)*	*(0.605)*	*(0.168)*	*(0.335)*	*(0.360)*	*(0.094)*	*(0.375)*	*(0.293)*	*(0.000)*	*(0.000)*
Vice Chair	2.27	12.59	7.07	-20.40	-61.55	-87.95	-22.58	-9.67	-14.16	-13.84	2.75	8.39
	(0.013)	*(0.651)*	*(0.197)*	*(0.336)*	*(0.309)*	*(0.306)*	*(0.037)*	*(0.143)*	*(0.835)*	*(0.114)*	*(0.121)*	*(0.263)*
	(0.001)	*(0.003)*	*(0.049)*	*(0.419)*	*(0.291)*	*(0.338)*	*(0.100)*	*(0.110)*	*(0.498)*	*(0.219)*	*(0.000)*	*(0.000)*
Governors	-7.68	10.61	5.10	-25.42	-86.36	-133.31	-20.14	-14.53	-26.24	-26.13	-3.67	0.97
	(0.135)	*(0.660)*	*(0.108)*	*(0.210)*	*(0.199)*	*(0.204)*	*(0.006)*	*(0.124)*	*(0.680)*	*(0.124)*	*(0.379)*	*(0.933)*
	(0.012)	*(0.036)*	*(0.026)*	*(0.368)*	*(0.238)*	*(0.293)*	*(0.114)*	*(0.248)*	*(0.706)*	*(0.259)*	*(0.000)*	*(0.000)*
FOMC	-5.43	8.89	-2.20	-63.34	-150.57	-161.09	-25.32	-38.07	-34.20	-56.84	-4.50	-17.67
	(0.283)	*(0.625)*	*(0.069)*	*(0.157)*	*(0.175)*	*(0.233)*	*(0.036)*	*(0.103)*	*(0.715)*	*(0.069)*	*(0.778)*	*(0.443)*
	(0.016)	*(0.008)*	*(0.044)*	*(0.327)*	*(0.306)*	*(0.309)*	*(0.061)*	*(0.255)*	*(0.660)*	*(0.186)*	*(0.007)*	*(0.218)*
Not FOMC	-31.11	-3.15	-3.55	-85.27	-223.71	-285.29	-31.29	-48.61	-44.63	-64.68	-13.26	-14.80
	(0.857)	*(0.248)*	*(0.005)*	*(0.101)*	*(0.109)*	*(0.165)*	*(0.013)*	*(0.091)*	*(0.438)*	*(0.114)*	*(0.683)*	*(0.463)*
	(0.045)	*(0.002)*	*(0.108)*	*(0.258)*	*(0.264)*	*(0.268)*	*(0.023)*	*(0.185)*	*(0.473)*	*(0.199)*	*(0.004)*	*(0.081)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month, 3-month and 6-month ahead forecasts over 2013M01–2020M02.

The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis.

Table 7: Forecast Evaluation: Text-augmented versus benchmark VAR - LDA on Speeches

	1M	3M	6M
Statements	7.95	8.94	6.86
	(0.001)	*(0.000)*	*(0.000)*
	(0.002)	*(0.000)*	*(0.000)*
Minutes	10.56	12.12	10.38
	(0.002)	*(0.000)*	*(0.000)*
	(0.007)	*(0.000)*	*(0.000)*
Chair	9.00	12.64	10.09
	(0.167)	*(0.006)*	*(0.009)*
	(0.438)	*(0.006)*	*(0.000)*
Vice Chair	2.04	10.07	5.81
	(0.821)	*(0.151)*	*(0.302)*
	(0.438)	*(0.006)*	*(0.000)*
Governor	-3.60	11.20	4.88
	(0.781)	*(0.140)*	*(0.470)*
	(0.390)	*(0.037)*	*(0.000)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month, 3-month and 6-month ahead forecasts for the Fed Funds Rate over 2013M01–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis.