Australian National University
Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

The Financial Instability – Monetary Policy Nexus: Evidence from the FOMC Minutes

CAMA Working Paper 25/2025
April 2025 (Revised in September 2026)

Dimitrios Kanelis
Deutsche Bundesbank

Lars H. Kranzmann
University of Münster

Pierre L. Siklos
Wilfrid Laurier University
Centre for Applied Macroeconomic Analysis, ANU

Abstract

We analyze how financial stability considerations expressed in Federal Open Market Committee (FOMC) meetings are associated with U.S. monetary policy implementation and communication. Utilizing natural language processing and language modeling to analyze FOMC minutes, we measure the prevalence and sentiment of mandate-related and financial stability-related discussions. Financial stability discussions increase after the Global Financial Crisis, especially during tightening cycles and the COVID-19 pandemic. Outside the zero lower bound (ZLB), greater attention to financial stability is associated with lower policy rates while during the ZLB, a tighter unconventional monetary policy stance is observed. Financial stability discussions are not significantly associated with sentiment in subsequent press conferences. Methodologically, we introduce a labeled dataset and apply explainable AI to improve the transparency of our text analysis.

Keywords

explainable artificial intelligence, financial stability, FOMC deliberations, monetary policy communication, natural language processing

JEL Classification

E44, E52, E58

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

The Centre for Applied Macroeconomic Analysis in the Crawford School of Public Policy has been established to build strong links between professional macroeconomists. It provides a forum for quality macroeconomic research and discussion of policy issues between academia, government and the private sector.

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

The Financial Instability – Monetary Policy Nexus: Evidence from the FOMC Minutes

Dimitrios Kanelis[1], Lars H. Kranzmann[2], Pierre L. Siklos[3]

September 2026

Abstract

We analyze how financial stability considerations expressed in Federal Open Market Committee (FOMC) meetings are associated with U.S. monetary policy implementation and communication. Utilizing natural language processing and language modeling to analyze FOMC minutes, we measure the prevalence and sentiment of mandate-related and financial stability-related discussions. Financial stability discussions increase after the Global Financial Crisis, especially during tightening cycles and the COVID-19 pandemic. Outside the zero lower bound (ZLB), greater attention to financial stability is associated with lower policy rates while during the ZLB, a tighter unconventional monetary policy stance is observed. Financial stability discussions are not significantly associated with sentiment in subsequent press conferences. Methodologically, we introduce a labeled dataset and apply explainable AI to improve the transparency of our text analysis.

JEL Classification: E44, E52, E58

Keywords: Explainable Artificial Intelligence, Financial Stability, FOMC Deliberations, Monetary Policy Communication, Natural Language Processing

[1]Deutsche Bundesbank, Wilhelm-Epstein-Str. 14, 60431 Frankfurt am Main, Germany

[2]Corresponding author: Department of Economics, University of Münster, Am Stadtgraben 9, 48143 Münster, Germany. E-mail: lars.kranzmann@wiwi.uni-muenster.de

[3]Department of Economics, Wilfrid Laurier University, Balsillie School of International Affairs, and the Centre for Applied Macroeconomic Analysis (CAMA)

¶We express our gratitude to Martin Bohl, Christian Groß, Pablo Guerron-Quintana, Tim Hagenhoff, Svatopluk Kapounek, Daria Kazakova, Malte Knüppel, Aloys Prinz, Peter Tillmann, and Eric Tong for their valuable comments and feedback. We are also grateful to the editor, Klodiana Istrefi, and two anonymous referees for helpful suggestions. We thank the participants at the 1st Piraeus International Economic Conference 2026, the 20th International Symposium on Econometric Theory and Applications at the University of Tokyo, the ECONDAT 2025 Spring Meeting by the Bank of England and King's College London, the 29th and 28th International Conferences for Macroeconomic Analysis and International Finance, the Conference on Real-Time Data Analysis, Methods, and Applications 2025 at the Czech National Bank, the 2025 Annual Conference of the Society for Nonlinear Dynamics and Econometrics, and the Rutgers University Celebration of Economic History in May 2025. We also extend special thanks to the participants of internal research seminars at the Deutsche Bundesbank and the University of Münster for their insightful and constructive feedback. We are grateful to Niklas Humann, Idir Chader, and Kessalia Süllau for their excellent research assistance. Part of the research for the paper was conducted while the second author was a visiting scholar at the Balsillie School of International Affairs. He is grateful for the hospitality. This paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or the Eurosystem.

1. Introduction

"Clean the mess afterward" is a famous expression referring to a policy that rejects preemptive central bank intervention in advance of financial imbalances or increasing financial risks (Bernanke and Gertler, 2001; Stein, 2013). If a boom ultimately turns out to be unsustainable, the central bank can then intervene and stimulate the economy. Following the Great Financial Crisis (GFC), this position earned its fair share of controversy between proponents who consider monetary policy an inefficient crisis prevention tool and critics who argue against crisis-induced welfare losses by leaning against the wind (Svensson, 2017). This shift in thinking has intensified scholarly efforts to explore how monetary policy interacts with financial stability. Related work by Boyarchenko et al. (2022) and Kashyap and Stein (2023) has intensified the debate concerning the financial stability risks posed by monetary policy.

A related policy debate concerns how risks to financial stability are reflected in the behavior of central banks and their reaction functions. Kashyap and Stein (2023) show, in a simple theoretical model, that if a central bank anticipates that monetary easing will increase financial instability in the medium term, it will behave differently than it would in the absence of such risks. Recent research demonstrates that the stance of monetary policy likely affects risk taking (Bauer et al., 2023) and that a policy rate that is too low for too long increases the probability of macro-financial instability (Grimm et al., 2026). Despite the growing interest in a potential trade-off between monetary policy and financial stability, there is little explicit evidence on how the balancing act between them shapes a central bank's monetary policy reaction function. This is true for the Federal Open Market Committee (FOMC), where the role of financial stability considerations in monetary policy decisions remains insufficiently understood. This gap is particularly relevant for understanding the zero lower bound (ZLB) period, during which unconventional monetary policy instruments, including large-scale asset purchases, became central to the implementation of monetary policy.

This study aims to close the above-mentioned gap in our understanding of monetary policy decision-making by answering three research questions: First, do financial stability considerations in FOMC deliberations help explain monetary policy decisions beyond the conventional determinants of a Taylor rule, and does their prevalence or sentiment matter? Second, does the relationship between financial stability considerations and monetary policy change between conventional and unconventional monetary policy regimes? Third, are financial stability considerations reflected with equal prominence in the Fed Chair's subsequent press conference communication?

To address these questions, we use the minutes of FOMC meetings, as they provide a systematic and comparable summary over time of the deliberations of the Committee as a whole. Since minutes are not limited to the views of individual policymakers, but also document the information, assessments, and policy options discussed at the committee level, they are a particularly valuable source of information. We are able to directly relate financial stability considerations discussed at each meeting to the corresponding monetary policy decision and examine whether the minutes provide additional information about the monetary policy stance beyond conventional macroeconomic determinants. At the same time, analyzing the process from FOMC deliberations to the policy decision, and the subsequent press conference, allows us to assess whether the same considerations are also reflected in the Fed Chair's public communication. Finally, the consistent structure of the minutes over time enables us to construct comparable meeting-level measures of the prevalence and sentiment of different topics.

Our empirical strategy consists of two steps. First, we use structural topic modeling (Roberts et al., 2016) to measure the prevalence of discussions related to the Fed's mandate objectives and financial stability. We then apply a fine-tuned RoBERTa model (Liu et al., 2019) to measure the sentiment expressed in these discussions. This allows us to distinguish between how prominently a topic is discussed and the tone with which it is discussed. We generate separate meeting-level indicators of topic prevalence and sentiment for discussions related to the mandate objectives and financial stability.

Second, we combine these text-based measures with real-time macroeconomic information and incorporate them into forward-looking Taylor rules (Coibion and Gorodnichenko, 2012). We examine whether the content of FOMC deliberations provides information about monetary policy decisions beyond conventional macroeconomic determinants and whether these relationships differ outside and during the ZLB period. We then use sentiment regressions to examine whether the same information from the deliberations helps explain variations in the sentiment of the Fed Chair's communication during the subsequent press conference. We further complement the quantitative results with narrative evidence from the FOMC transcripts.

Our study makes several contributions to the literature. First, we provide new evidence that the relationship between financial stability considerations and monetary policy depends on the monetary policy regime. Outside the ZLB, an increase in financial stability discussions is associated with a lower policy rate, consistent with the evidence from public speeches

documented by Istrefi et al. (2023). During the ZLB, however, this relationship reverses. Using shadow rate estimates of Wu and Xia (2016) and Krippner (2013), we find that more discussion of financial stability issues is associated with a tighter monetary policy stance. Importantly, this result is driven by the extent to which financial stability is discussed rather than by the sentiment expressed in these discussions. Our findings therefore suggest that the policy implications of financial stability considerations are regime dependent. While financial stress can be associated with monetary easing, concerns about financial vulnerabilities can also become relevant for the withdrawal of accommodation when interest rates remain low for longer. This interpretation is consistent with the mechanism emphasized by Kashyap and Stein (2023). Narrative evidence from FOMC transcripts provides further support for our interpretation, showing that participants explicitly discussed reach-for-yield behavior and the potential build-up of financial imbalances as side effects of prolonged monetary easing.

Second, we examine whether the information contained in FOMC deliberations is also reflected in the Chair's subsequent monetary policy communication. More positive discussions of the mandate objectives are associated with higher policy rates and more positive sentiment in the Chair's subsequent introductory statement. For financial stability, we find a different pattern. Financial stability discussions are associated with the monetary policy stance, but not with the sentiment of the Chair's subsequent press conference communication. This suggests that financial stability considerations relevant for policy decisions are not necessarily communicated with the same degree of prominence. FOMC minutes can, therefore, provide additional information about considerations underlying monetary policy decisions.

Third, our study contributes methodologically to the measurement of central bank deliberations. We fine-tune a RoBERTa language model (Liu et al., 2019) using a manually labeled dataset tailored to the language of the FOMC minutes. This provides a unified framework for measuring sentiment across different discussion topics without relying on separate topic-specific dictionaries. To address the limited transparency of transformer-based language models relative to dictionary approaches, we employ SHapley Additive exPlanations (SHAP) values (Shapley, 1953; Lundberg and Lee, 2017) to make the model's classifications interpretable. We also derive a SHAP-based sentiment dictionary as a robustness check and as an alternative to widely used existing dictionaries (e.g., Loughran and McDonald (2011) and Correa et al. (2021)). As a consequence, we make available a novel dataset of annotated sentences from FOMC minutes, thereby providing a reusable resource for fine-tuning or for prompting large language models in future research.

The remainder of the paper is structured as follows. Section 2 reviews the related literature and highlights the studies most relevant to our analysis. Section 3 describes the textual data, machine learning models, and the labeled dataset developed for this study. Section 4 presents the empirical framework, followed by the results and extensive robustness checks in Section 5. The final section concludes.

2. Literature Review

Monetary policy and financial stability are closely intertwined, allowing the central bank to have "a natural role to play in ensuring financial stability" (Schinasi, 2003). Indeed, as Adrian and Liang (2018) point out, recent research has emphasized a link between accommodative monetary policy and an increase in risks to financial stability. This raises the question: to what extent should financial stability considerations be embedded within the central bank's policy objectives (Smets, 2014)? This is especially true when communicating about financial stability (Born et al., 2014) since this instrument allows central banks to preempt financial instabilities through warnings (Svensson, 2003) or, conversely, provoke market disturbances (Cukierman, 2009).

To address our research questions, we draw on two related strands of literature. The first examines minutes and transcripts of deliberations to understand the reasoning behind FOMC monetary policy decisions. A seminal study by Hansen et al. (2018) utilizes the natural experiment created by the 1993 release of all FOMC transcripts. Since FOMC members initially expected these deliberations to remain private, the study analyzes how this newfound transparency influenced their behavior. As part of their research design, Hansen et al. (2018) apply Latent Dirichlet Allocation (LDA), a topic modeling approach of Blei et al. (2003), to identify text sections containing information on the policy preferences of the FOMC members. The authors find that transparency amplifies the effect of career concerns. Shapiro and Wilson (2022) apply sentiment analysis to the transcripts to determine the implicit inflation target of the FOMC for the period 2000 until 2011. The authors conclude that the implicit inflation target was lower than the publicly assumed target of 2%.[1] The central insight of this literature strand is that FOMC deliberations document information about US monetary policy beyond the information that deviations of macroeconomic variables from target values provide.

[1]The Fed introduced a medium-term inflation target of 2% in 2012.

The second strand of the literature examines various central bank publications to determine how central bankers prioritize financial stability. Human-based or automated sentiment analysis techniques are employed to measure optimism or pessimism. The resulting indicators are often added to standard inflation and output or unemployment gaps found in Taylor rules. For example, Peek et al. (2016) analyze FOMC meeting transcripts before the GFC and find that the frequent use of financial instability-related terms indicates a significant impact of financial stability considerations on the Fed's policy decisions. Similarly, Oet and Lyytinen (2017) classify FOMC minutes into topics, demonstrating that incorporating a broader range of discussed topics, especially financial stability, into a Taylor rule improves our understanding of the Fed's monetary policy. Wischnewsky et al. (2021) extract textual information from Humphrey-Hawkins hearings transcripts to integrate financial stability sentiment into a Taylor rule. Their findings suggest that negative sentiment towards financial stability correlates with a more accommodative monetary policy, thereby improving the Taylor rule's explanatory power.

Istrefi et al. (2023) identify the financial stability content of Fed speeches using LDA and measure the negativity of the tone based on the financial stability dictionary by Correa et al. (2021). Among these studies, Istrefi et al. (2023) is the study closest to ours because they also combine text-based measures of financial stability communication with a forward-looking Taylor rule for the Federal Reserve. Their findings provide an important benchmark for our analysis, as they show that greater attention to financial stability is associated with a more accommodative monetary policy before the GFC, while this relationship is no longer present after the crisis. In a recent study, Benchimol et al. (2025) investigate the characteristics of the Federal Reserve's communication during the COVID-19 crisis. Fed communication was found to be notably more reactive during the pandemic compared to other crises, as evidenced by differences in content, sentiment, and timing.

More recent studies leverage advances in language modeling. Kanelis and Siklos (2025) use a financial adaptation of the widely used Bidirectional Encoder Representations from Transformers (BERT) model to measure sentiment in the ECB Executive Board's price and financial stability-related speeches. The authors find that inter-meeting speeches on monetary policy and inflation can explain changes in the sentiment of the introductory statements of the next press conferences. In contrast, financial stability speeches offer no additional explainability. Bertsch et al. (2025) examine how senior Fed officials interpret the dual mandate. Using BERT and RoBERTa to analyze speeches, they find that financial stability is the most frequently discussed topic outside the dual mandate. Financial stability

considerations are only considered in the context of fulfilling the Fed's mandate, and not seen as necessarily triggering a trade-off with the central bank's dual mandate.

3. Topic and Sentiment Analysis Applied to FOMC Minutes

3.1. Textual Data

To understand how the FOMC discusses financial stability in relation to its mandate, we utilize the minutes of the meetings. We opted to use the minutes instead of the transcripts since the minutes reflect the position of the entire FOMC and are less susceptible to noise from the opinions of individual participants. This characteristic is particularly useful because individual views, even when strongly held, do not necessarily translate into the Committee's final policy decision (Kazinnik and Sinclair, 2026). Furthermore, the minutes offer a detailed summary of the deliberations while still capturing the range of perspectives among the Board of Governors and Reserve Bank presidents on monetary policy (Meade et al., 2015). Using the minutes ensures consistency in content, language, and document structure across our corpus. As highlighted by Danker and Luecke (2005), the uniformity in language and structure aids in understanding policy communication. Compared to the more detailed transcripts of FOMC meetings, which the Fed releases with a five-year delay, the minutes have been published three weeks after each meeting since December 2004, whereas previously they were released after the subsequent meeting. It is therefore understandable why markets and the public are interested in the minutes since the information contained is less stale than in the transcripts.[2] Concurrently, the minutes have substantially increased in length, suggesting a potential enhancement in the degree of detail provided (Meade et al., 2015). Our FOMC meeting minutes corpus covers January 1993 to December 2022. We choose this period since, starting in 1993, the "Record of Policy Actions" was combined with the "Minutes of Actions", resulting in the minutes in their current form (Danker and Luecke, 2005).

The second text corpus in our analysis consists of the Fed Chair's introductory statements at press conferences and the Chair's responses during Q&A sessions. The first press conference took place in April 2011, following an FOMC meeting. During our study period, press conferences were not held after every meeting. Until the end of 2018, they were conducted

[2]That said, Acosta (2023) demonstrates that the similarity between the minutes and the transcripts has grown over time, reaching a relatively high level during our study period. This is not surprising because participants are aware that transcripts will eventually be published.

only with the publication of the Summary of Economic Projections (SEP). Since then, the Fed has held press conferences after every meeting. Our corpus comprises the press conferences from April 2011 to December 2022 consisting of introductory statements and Q&A sessions.[3]

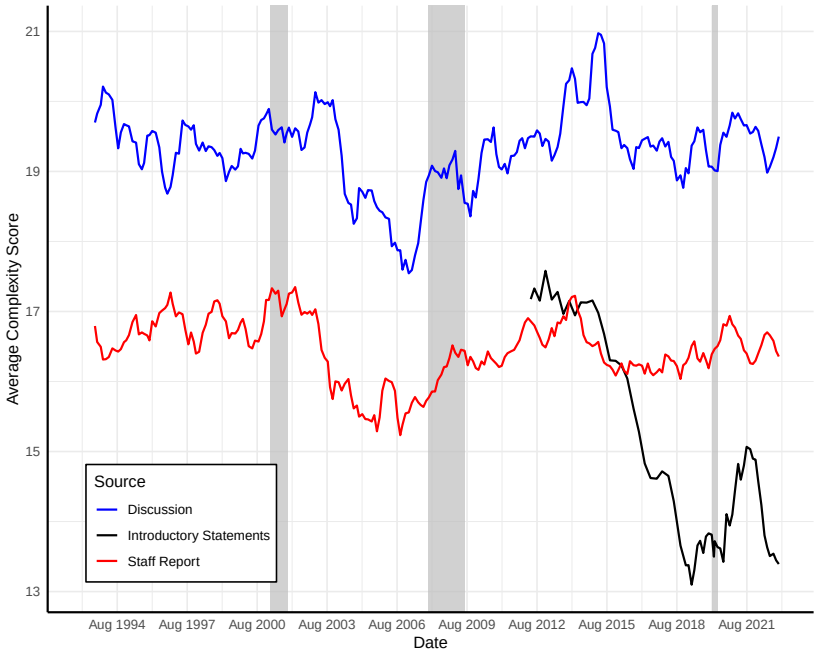
When applying NLP to documents like the FOMC minutes, it is important to note that the language used can be more complex than everyday language or standard economics and financial communication (Hernández-Murillo and Shell, 2014). To illustrate this complexity, we calculate an average complexity score based on four different readability measures: Automated Readability Index, Gunning-Fog, Coleman-Liau, and Flesch-Kincaid (Siklos et al., 2018). These measures aim to capture text complexity as it relates to the US educational system. We divided the minutes into the staff report and the discussion sections. Figure 1 shows that the complexity of the staff reports and discussions develops similarly in terms of textual complexity, with the discussions inside the FOMC being the most complex. Note also that the introductory statements have a significantly lower complexity score, and the Staff Report is always less complex than the committee discussions. According to the readability measures, discussions effectively can only be understood by university graduates. This motivates our methodological contribution of fine-tuning an LLM to handle the sophisticated protocol language of the minutes that we introduce in section 3.3.

3.2. Topic Identification and Analysis

To identify the topics of interest, we focus on the sections entitled *Members' View of the Economic Situation* and *Policy Decision*. These sections summarize the Committee's assessment of economic and financial developments and the deliberations underlying the policy decision. Since our objective is to measure the content of these deliberations, the choice of the textual unit is important. We use paragraphs as the input documents for the topic identification algorithm because the FOMC minutes are purposefully structured so that individual paragraphs typically discuss one specific issue or policy consideration. This paragraph-level structure is particularly useful in our setting, as it reduces the risk of mixing several economically distinct concepts within a single textual unit. Few studies, such as Jegadeesh and Wu (2017) and Oet and Lyytinen (2017), exploit this characteristic of the minutes. However, they do not perform automated classification, even though the document structure is well

[3]We obtained all textual data directly from the Federal Reserve Board of Governors website.

Figure 1: Complexity Measures for FOMC Documents 1993-2022



Note: The figure shows the average textual complexity of staff report and discussion sections of the FOMC minutes from January 1993 to December 2022. It also reports the complexity of the available introductory statements from press conferences, which began in April 2011. The complexity score is calculated using the mean of the Automated Readability Index, Gunning-Fog Score, Coleman-Liau Score, and Flesch-Kincaid Grade Scores. NBER-dated recession periods, as defined by NBER, are indicated by shaded areas.

suited for this purpose.[4]

We use the Structural Topic Model (STM) of Roberts et al. (2016) to categorize the individual paragraphs of the minutes into latent discussion topics. The STM has become an established unsupervised learning approach in economics because it recovers systematic themes from large text corpora while retaining an interpretable representation of topics through their word distributions (Dybowski and Kempa (2020), Ferrara et al. (2022), Bohl et al. (2023)). For this purpose, we perform the following preprocessing steps on the individual paragraphs: converting all words to lowercase, removing numbers and stop words, and applying word stemming.

A central modeling choice is the number of topics, K. Following Roberts et al. (2019), we select K by considering the trade-off between semantic coherence (Mimno et al., 2011) and topic exclusivity (Bischof and Airoldi, 2012). Topic models with a larger number of topics typically generate more exclusive and differentiated topics, whereas models with fewer topics tend to increase semantic coherence by combining related discussions into broader themes. For our baseline specification, we use $K = 21$, which provides a suitable balance between these criteria and yields topics that can be mapped into economically meaningful categories.

We do not impose a small number of topics for reasons of parsimony alone. Several concepts that are central to our analysis are closely related but economically distinct, including financial vulnerabilities, financial conditions, housing and mortgage markets, monetary policy implementation, labor market conditions, inflation, and the broader economic outlook. A low-dimensional topic model may therefore generate topics with high semantic coherence while simultaneously combining conceptually distinct discussions into broad macroeconomic or financial themes. This semantic coherence would reduce our ability to isolate the specific concepts that are central to our analysis.[5] The technical details of the topic-selection procedure, including the semantic coherence–exclusivity frontier and topic identification and aggregation, are in Appendix A.

After estimating the STM, we assign topic labels using both the highest-probability word stems and representative high-probability paragraphs. This step is important because the

[4]Topic modeling performs best when the input documents are relatively monothematic, as the algorithm assigns documents to topics based on the underlying word distributions.

[5]To illustrate this point, Appendix A.2 presents the topics obtained from a highly parsimonious specification with $K = 5$. These topics exhibit high semantic coherence but they merge several economically distinct discussions that are separated in our baseline specification.

economic interpretation of a topic cannot rely on word lists alone. Table 1 reports the complete list of topic labels together with their most influential word stems.[6] We then map the individual STM topics into two broader discussion categories: *Mandate* and *Financial Stability*. We follow Moschella and Pinto (2019) by first identifying topics using a data-driven topic model and then aggregating them ex post to align the resulting measures with the focus of the analysis. This sequential procedure follows established practice in the topic-modeling literature (e.g., Dyer et al., 2017; Istrefi et al., 2023).

The *Mandate* aggregate captures paragraphs related to the Federal Reserve's dual mandate and the economic conditions against which the Committee evaluates progress toward its objectives. We combine the topics *Business Investment and Production*, *Labor Market Conditions*, *Dual-Mandate Objectives*, *Inflation Expectations*, *Real Activity and Demand*, *Economic Risks and Shocks*, *Policy Implementation*, *Policy Stance and Risks*, *Inflation Outlook*, and the *Growth Outlook and Risks*.[7]

The *Financial Stability* aggregate captures paragraphs from the topics *Financial Conditions*, *Financial Vulnerabilities and Risk*, and *Housing and Mortgages*. We adopt this broader definition because financial stability concerns in FOMC deliberations often arise from discussions of credit conditions, liquidity, market functioning, mortgage markets, and housing-related vulnerabilities. While financial conditions are not identical to financial stability, the two are closely connected in the conduct and transmission of monetary policy (Adrian and Liang, 2018; Kashyap and Siegert, 2020). Housing-related discussions are included because housing market developments played a central role in the build-up and transmission of financial vulnerabilities during our sample period. This aggregation is consistent with the multidimensional view of financial stability communication in Istrefi et al. (2023) and with the view that monetary policy can affect risk-taking and vulnerabilities and "gets into all the cracks" of the financial system (Stein, 2013).

Having constructed these two aggregate discussion categories, we calculate topic proportions at the meeting level using equation (1).

---

[6]Representative paragraphs for the topics used in the construction of our main variables are provided in Appendix A.1.

[7]In selecting mandate-related topics, we also consider broader discussions of output and economic activity, since current and expected unemployment developments are closely linked to real activity. During her tenure, Fed Chair Yellen (2013) stated that "[t]he general development of output is also considered when assessing unemployment indicators." In Appendix B.1, we show that our results and conclusions are robust when using a narrow definition restricted to *Labor Market Conditions*, *Dual-Mandate Objectives*, *Inflation Expectations*, and *Inflation Outlook*.

Table 1: Topic Identification Results

Procedural Information	**Financial Conditions**^{FS}	**Financial Vulnerabilities and Risk**^{FS}	**Business Investment and Production**^{M}	**Labor Market Conditions**^{M}	**Dual-Mandate Objectives**^{M}	**Household Sector**
releas	secur	financi	busi	labor	particip	consum
statement	feder	market	invest	particip	committe	spend
vote	reserv	particip	capit	market	polici	household
short	agenc	credit	report	rate	rate	incom
approv	hold	bank	spend	unemploy	inflat	recent
encompass	committe	risk	firm	employ	monetari	growth
p.m	treasuri	condit	contact	wage	goal	increas
meet	polici	note	equip	worker	feder	effect
inclus	agre	term	product	job	guidanc	support
press	term	increas	continu	increas	percent	confid
sentenc	support	loan	industri	condit	time	fiscal
paragraph	market	yield	construct	growth	run	sale
text	level	equiti	increas	forc	employ	gain
percentag	maintain	rate	profit	continu	maximum	factor
unanim	condit	remain	expenditur	remain	achiev	retail

Foreign Economic Activity	**Housing and Mortgages**^{FS}	**Inflation Expectations**^{M}	**Voting Procedures**	**Real Activity and Demand**^{M}	**Policy Communication**	**Credit Easing**
u.	hous	inflat	direct	growth	committe	purchas
foreign	sector	expect	reserv	expans	statement	asset
export	particip	particip	feder	econom	meet	committe
economi	activ	term	polici	economi	polici	pace
econom	price	measur	committe	inventori	agre	program
growth	report	run	conclus	activ	econom	particip
domest	mortgag	percent	bank	demand	discuss	market
trade	home	market	domest	prospect	languag	econom
dollar	continu	committe	instruct	current	decis	outlook
demand	construct	month	author	develop	chang	progress
effect	contact	remain	accord	quarter	futur	addit
global	rate	continu	execut	moder	risk	condit
abroad	district	recent	transact	rate	postmeet	continu
develop	sale	price	york	outlook	announc	reduct
market	recent	object	discuss	busi	communic	improv

Money and Credit Aggregates	**Economic Risks and Shocks**^{M}	**Outlook Assessment**	**Policy Implementation**^{M}	**Policy Stance and Risks**^{M}	**Inflation Outlook**^{M}	**Growth Outlook and Risks**^{M}
rang	particip	project	rate	polici	inflat	econom
growth	econom	rate	fund	tighten	price	growth
committe	risk	gdp	feder	risk	increas	particip
reserv	activ	real	inflat	economi	pressur	risk
monetari	note	forecast	rang	committe	cost	outlook
econom	suppli	percent	committe	monetari	labor	pace
aggreg	economi	growth	target	action	product	continu
period	polici	run	condit	period	rise	expect
price	pandem	unemploy	outlook	economi	expect	meet
debt	continu	polici	polici	move	resourc	recent
condit	outlook	particip	expect	eas	energi	activ
consist	support	feder	percent	view	recent	downsid
accept	remain	assess	market	stanc	core	quarter
slight	fiscal	inflat	assess	intermeet	remain	moder
intermeet	uncertainti	monetari	direct	direct	market	anticip

Note: The table reports the 15 most influential word stems for each STM topic, ranked by the probabilities that the word appears in the respective topic. Topic labels are assigned based on dominant word stems and representative paragraphs (see Appendix A.1). Superscripts indicate topics included in the aggregate measures: M denotes *Mandate* and FS denotes *Financial Stability*.

$$TP_t^T = \frac{\#Sentences_t^{D_T}}{\#Sentences_t^{All}} \tag{1}$$

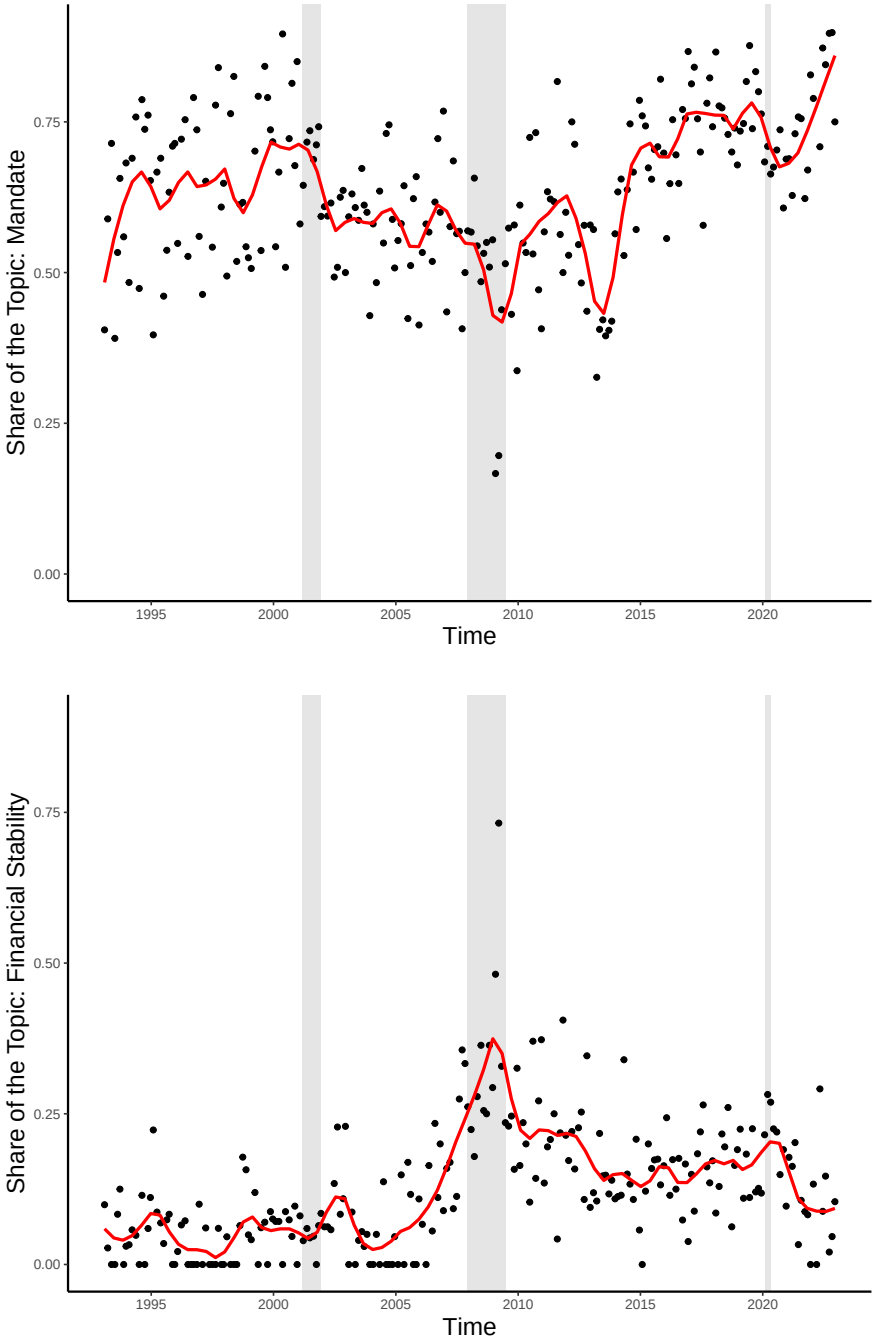
TP_t^T, for $T \in \{$*Mandate, Financial Stability*$\}$, on a given meeting date t is calculated by dividing the total number of sentences in paragraphs labeled with topic T by the total number of sentences in the same document. The numerator represents the sum of sentences from paragraphs associated with topic T, where D_T represents the set of topic IDs linked to T. The denominator represents the total number of sentences from all paragraphs discussed on meeting date t. Figure (2) illustrates the progression of topic proportions over time.

Figure (2) visualizes the topic proportions for *Mandate* (upper panel) and *Financial Stability* (lower panel) from January 1993 until December 2022. The dots represent the topic proportions in the minutes of each meeting. To describe the trend, we add a LOESS curve to the individual meeting estimates. Visual inspection allows us to derive some stylized facts about how the FOMC discusses these topics. First, discussions on *Mandate* account for a larger share of the minutes than discussions on *Financial Stability* in most meetings. Second, the correlation between the two topic proportions is -0.41 ($p < 0.01$), suggesting that the relative attention devoted to mandate-related and financial stability-related discussions varies substantially across meetings.

An exception happened during the GFC, when the *Financial Stability* topic proportion temporarily exceeded the *Mandate* topic proportion primarily due to a high focus on the housing market, mortgage defaults, and broader financial market disruptions. Since the GFC, the *Mandate* topic proportion has again exceeded the *Financial Stability* topic proportion and gained greater importance during FOMC meetings. Discussions about unemployment are a primary reason for this change. Before the GFC, the FOMC viewed unemployment partly as a structural issue related to globalization and international trade. Following the financial crisis, the FOMC assigned monetary policy a more prominent role in reducing unemployment, even after taking the output gap into account. This observation is consistent with the results of Bohl et al. (2023), who analyzed the Board of Governors' communication on unemployment in speeches.

The lower panel of figure (2) illustrates the significant rise in the proportion of discussions on *Financial Stability* in FOMC minutes since the GFC. In contrast, during the 1990s and early 2000s, financial stability was only of minor relevance in many meetings. A minor exception occurred during the financial turmoil in the early 2000s, particularly around the

Figure 2: Topic Proportions of Mandate-Related and Financial Stability Discussions in FOMC Minutes (1993–2022)



Note: The upper figure displays the proportion of mandate-related topics identified by the STM in the FOMC meeting minutes from January 1993 to December 2022, while the lower figure shows the proportion of financial stability-related topics over the same period. Both are based on meeting frequency. The red line represents a locally estimated scatterplot smoothing (LOESS) curve, illustrating the underlying trend in the data. NBER-dated recession periods are indicated by shaded areas.

dot-com bust. However, the proportion of discussions on financial stability reached a local low in the years between the dot-com bust and the housing crisis that triggered the GFC.[8] The GFC marked a structural shift in discussion intensity and led to a more persistent integration of the topic into FOMC deliberations. Since 2011, the proportion of discussion on *Financial Stability* in the minutes has remained around 20%, with the highest increase occurring during the COVID-19 pandemic.

3.3. Developing a Federal Reserve Minutes Sentiment Classifier

Identifying mandate-related topics omits the possibility that sentiment also plays a role in the conduct of monetary policy. Consistent with the literature (Bohl et al. (2023), Istrefi et al. (2023)), we implement sentiment analysis only after categorizing the individual topics to ensure that the sentiment indicators are not biased by content without any connections to either discussions on topics related to the *Mandate* or to the *Financial Stability* aggregates. For the sentiment analysis, we choose the Positive-Negative metric since this metric offers straightforward information about how the FOMC frames its narratives and allows comparison between different topics.[9] Furthermore, we resort to language modeling for the sentiment analysis instead of relying on dictionary methods because language models can consider the context of the text and are less susceptible to noise (Mishev et al. (2020), Huang et al. (2023)).

As described in section 3.1, the protocol language used in the minutes is formal and structured, differing significantly from the diverse language styles typically encountered by LLMs. To prevent the higher textual complexity and the unique protocol of central bank communication styles from biasing our analysis, we follow the literature on supervised textual machine learning and fine-tune a pre-trained model for this specific context. This brings us back to the question of the appropriate document level to focus on. Shifting our focus from paragraphs to individual sentences, we recognize that sentences are meaningful units that can easily be comprehended and evaluated by a human reader (Maibaum et al., 2024), and they serve as particularly effective units of observation in central bank documents (Kanelis and Siklos, 2025). In contrast, paragraphs often contain a mix of sentiments, making them

[8]This pattern is consistent with the broader observation that financial vulnerabilities are often underappreciated before crises become visible, see Reinhart and Rogoff (2009).

[9]An alternative would be to use a Hawk-Dove metric. Despite its relevance in the literature (Apel et al., 2022), this metric works best for discussions about inflation or unemployment but is less useful for topics like financial stability. A Positive-Negative metric is more intuitive for financial stability analyses (Correa et al., 2021).

harder for human readers to evaluate consistently and therefore less suitable as reliable units for sentiment classification. This leads to the creation of a novel dataset consisting of 2000 manually labeled sentences randomly chosen from the discussion section of the minutes.

To ensure accuracy and objectivity, each sentence was independently classified by the first two authors, with the third author reviewing the classifications and deciding in case of disagreement. The process involved systematically categorizing each sentence based on its overall sentiment tone from a central bank's perspective. Specifically, we interpreted positive sentiment as language that indicates that economic conditions are consistent with or approaching the objectives of the FOMC. Examples include inflation at or heading toward the inflation target, rising labor utilization toward full employment, stabilization of real output growth around potential, and smoothly functioning financial markets with limited volatility. In contrast, negative sentiment referred to statements suggesting that inflation is deviating from the target in either direction, that labor market weakness persists or overheating pressures are building, that growth is significantly below or at unsustainably high levels above potential, or that financial conditions are strained due to increased spreads, illiquidity, or emerging vulnerabilities.

Each sentence was evaluated in isolation, to create a versatile dataset that can be used in various applications. Therefore, directional statements such as "rising" or "falling" were only assessed as positive or negative if the context clearly indicated whether the development was bringing conditions closer to or further away from policy targets. We assessed the degree of agreement between the first two authors by using Cohen's kappa and calculated a value of 0.64, which signals substantial agreement (Cohen (1960), Landis and Koch (1977)).[10] The labeled dataset comprises a total of 793 neutral sentences, 648 negative sentences, and 559 sentences labeled as positive. The creation of this dataset is a notable contribution on its own, independent of our research questions, as it can be valuable for researchers looking to fine-tune or prompt LLMs for future research.[11]

We choose the pre-trained RoBERTa model for further analyses. Liu et al. (2019) introduce RoBERTa by building on the BERT model from Devlin et al. (2019). Pre-training of BERT is conducted unsupervised via two techniques in parallel: Masked Language Modeling (MLM),

[10]Cohen's kappa is a statistical measure that evaluates inter-rater reliability. Cohen's kappa has a value of 1 when there is perfect agreement, 0 when the agreement is equivalent to chance, and -1 when there is complete disagreement (Cohen, 1960).

[11]The dataset is publicly available in our Google Drive repository. Access is provided in Appendix D.

in which random words are masked, and the model predicts them based on the preceding and following words, and Next Sentence Prediction (NSP), which evaluates whether two sentences follow each other. It is trained on general-language corpora such as the English Wikipedia (Devlin et al., 2019). RoBERTa is a modification of BERT with important changes: it omits the NSP component, uses a larger dataset for training, and extends the sequence length of the sentences in the pre-training for improved model performance. Like BERT, RoBERTa is also trained in general language but performs better than BERT due to these changes (Liu et al., 2019). We conducted fine-tuning using 70% of our dataset for training, 15% for validation, and 15% for testing.[12]

The model evaluation is based on precision, recall, and the F1-score. We fine-tune both RoBERTa and FinBERT (Huang et al., 2023) on our manually labeled training data and select RoBERTa for the baseline analysis because it achieves the strongest classification performance on the held-out evaluation sample.[13] To assess the incremental contribution of the fine-tuned language model relative to established dictionary-based methods, in Appendix D, we compare its performance with both the dictionary by Loughran and McDonald (2011) and the financial stability dictionary developed by Correa et al. (2021). All three approaches are first evaluated on the same held-out sample of 300 manually labeled sentences from the FOMC minutes. The fine-tuned RoBERTa model achieves an overall accuracy of 0.79, compared with 0.54 for the LM dictionary and 0.53 for the dictionary by Correa et al. (2021), and obtains higher F1-scores in each sentiment category.

Because the dictionary by Correa et al. (2021) was specifically designed for financial stability related texts, we additionally conduct a more targeted comparison using a separate out-of-sample dataset of 300 manually labeled sentences drawn from financial stability discussions.[14] None of these sentences was used to fine-tune the RoBERTa model. Even in this domain-specific evaluation, the fine-tuned RoBERTa model outperforms the dictionary by Correa et al. (2021), achieving an accuracy of 0.78 compared with 0.62 and higher F1-scores for positive, neutral, and negative sentences. A methodological advantage of our approach is that it provides a unified framework for the entire analysis, with higher classification accuracy, rather than combining separate topic-specific dictionaries, as done in previous

[12]Appendix D provides a description of the fine-tuning procedure, including hyperparameter selection.

[13]This result is consistent with Pfeifer and Marohl (2023), who find that RoBERTa performs particularly well in classifying sentiment in central bank speeches.

[14]We make this evaluation dataset also publicly available (see Appendix D).

studies.

A common criticism of AI approaches for the classification of qualitative data is the opacity of the model. Consequently, and despite the significant advances in language modeling, researchers continue to use manually crafted lexicographic approaches (Cieslak and McMahon (2024), Aruoba and Drechsel (2026), Chadha et al. (2026)). However, this does not align with the current state of knowledge in machine learning (Ali et al., 2023). A proven approach to overcoming the opacity of AI models is the use of SHAP values. These are based on the game-theoretical concept introduced by Shapley (1953) and provide a means to make the contribution of each individual input variable to a model's prediction comprehensible. SHAP values allow for the quantification of a variable's influence on the model by considering all possible sequences in which the variable could be added to the model. This not only identifies the individual contribution of each variable to the prediction but also accounts for how this contribution is affected by the presence of other variables. SHAP provides both local explanations, which highlight the significance of variables for a specific prediction, and global explanations, which summarize the importance of the variables across all predictions.[15]

In our context, RoBERTa takes an input sentence and predicts the probability of different sentiment classes - in our case, positive, neutral, and negative sentiment. Each word in the sentence contributes differently to these predictions, and SHAP values help quantify the influence of each word. Each sentence has a base score that represents the model's expected prediction if the entire sentence were masked, with all words unknown. This base value serves as a reference point, indicating what the model would predict in the absence of any contextual information. SHAP values illustrate, in an additive manner, how revealing each word influences the model's output, shifting from the base value to the final predicted value (Lundberg and Lee, 2017). Following this procedure, Table (2) shows the most relevant words in the minutes that contribute to the sentiment classification of the FOMC minutes.[16]

With our model framework in place, we calculate the sentiment for the narratives on *Mandate* and *Financial Stability* using the following formula:

[15]SHAP values are used both in regression analyses in economics using machine learning (Beckmann et al., 2023) and for enhancing interpretability in automatic image recognition.

[16]In Appendix E, Figure E1 provides the results of the Shapley decomposition on a number of classification examples from the minutes during different time periods.

Table 2: Top 20 Positive and Negative Sentiment Tokens Based on SHAP Values

Positive Words	**Score**	**Negative Words**	**Score**
recede	0.5868	unacceptably	-0.5414
subside	0.4391	uncomfortably	-0.5319
bright	0.4237	worrisome	-0.4828
optimistic	0.3781	sluggishness	-0.4614
upbeat	0.3673	deteriorate	-0.4007
lessen	0.3613	concerned	-0.3912
resilient	0.3446	inconsistent	-0.3825
positive	0.3328	falter	-0.3786
optimism	0.3307	worried	-0.3648
beneficial	0.3254	uncertainty	-0.3566
improve	0.3188	uneven	-0.3491
buoy	0.3112	unsettle	-0.3407
mute	0.3073	weaken	-0.3381
check	0.2803	pessimistic	-0.3376
benign	0.2472	mislead	-0.3367
favorable	0.2424	subpar	-0.3325
favorably	0.2380	unsustainable	-0.3293
rebound	0.2372	depressed	-0.3263
encourage	0.2367	disappointing	-0.3259
abate	0.2354	hardship	-0.3172

Note: This table presents the top 20 words ranked by their SHAP scores, calculated using our fine-tuned RoBERTa model. Each word has a SHAP value for each sentiment category of the RoBERTa model. We calculate the score by measuring the difference between a word's contribution to positive sentiment and its contribution to negative sentiment across all its occurrences in the text corpus. Thus, the SHAP score is defined as $Score = PositiveSentiment - NegativeSentiment$. Words with the highest positive scores exert the greatest influence in classifying sentences as positive, while those with the highest negative scores most strongly impact negative sentiment classification. Words that appeared fewer than 10 times were excluded from the analysis.

$$Sentiment_t^T = \frac{\#Sentences_t^{Positive,T} - \#Sentences_t^{Negative,T}}{\#Sentences_t^{Positive,T} + \#Sentences_t^{Neutral,T} + \#Sentences_t^{Negative,T}} \quad (2)$$

We calculate the sentiment for each topic $T \in \{$*Mandate, Financial Stability*$\}$ discussed during the FOMC meeting on date t by taking the difference between the number of positive and negative sentences related to T and dividing it by the total number of sentences on that topic.[17] The sentiment series for both *Mandate* and *Financial Stability* are visualized in Figure (3).

Sentiment analysis provides further stylized facts about the FOMC deliberations. Figure 3 illustrates the evolution of sentiment on *Mandate* (upper panel) and *Financial Stability* (lower panel). The sentiment indicators for *Mandate* and *Financial Stability* have a correlation of 0.25 ($p < 0.01$) across the full sample. In the 1990s, sentiment regarding *Mandate* was generally balanced, with only a brief downturn during the early 2000s recession. Sentiment rose significantly in the mid-2000s but then plummeted during the GFC, indicating heightened pessimism. This trend gradually improved throughout the 2010s, only to sharply decline again with the onset of the COVID-19 pandemic and the rising inflation since 2021.

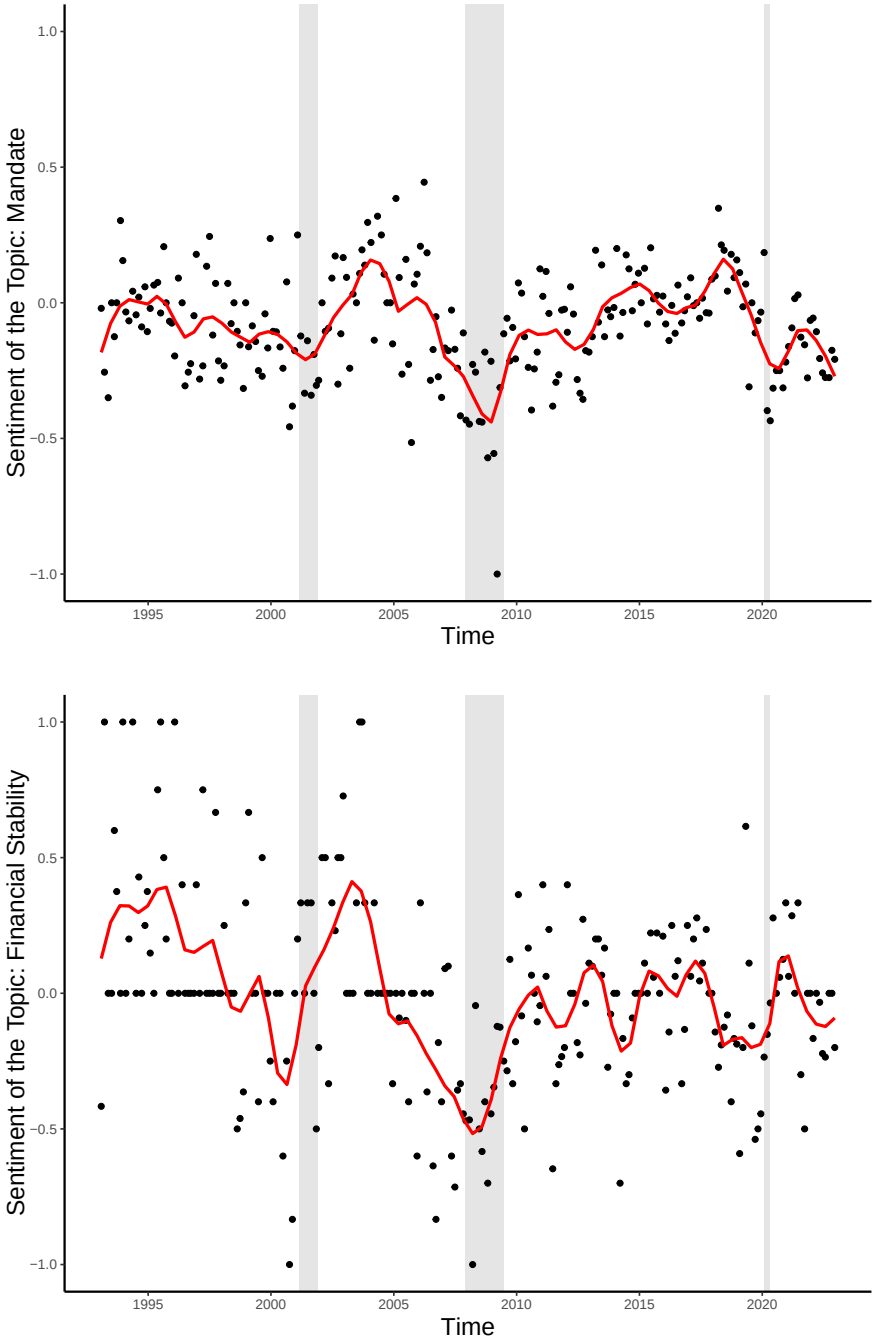
The lower panel presents the sentiment of discussions dealing with *Financial Stability*. Before the GFC, it is important to note that the overall sentiment value is influenced by a limited number of sentences due to the low or even non-existent proportion of this topic. This results in greater volatility compared to the period following the GFC, when financial stability became a regular discussion topic in meetings. As anticipated, we observe increased pessimism during the 2001 dot-com bubble and the 2008/2009 GFC, when the sentiment indicator hits an all-time low. Since then, discussions on financial stability have remained more balanced and gradually positive until the onset of the COVID-19 pandemic.

4. Econometric Framework

After quantifying the information in the minutes, we analyze how it helps us understand monetary policy decisions and communication. We do so by implementing a framework that simulates the sequence of the FOMC meeting from the deliberations until the publication of the decisions. The FOMC meeting process begins with extensive preparations, where

[17] We divide by the total number of sentences to account for variations in document length over time.

Figure 3: Sentiment Analysis of FOMC Discussions Using a Fine-Tuned RoBERTa Model (1993–2022)



Note: The upper figure displays sentiment analysis results for *Mandate* from January 1993 to December 2022, while the lower figure shows sentiment analysis for *Financial Stability* in the FOMC meeting minutes during the same period. Both analyses utilize a fine-tuned RoBERTa model applied at the sentence level. The red line represents a LOESS curve, illustrating the underlying trend in the data. NBER-dated recession periods are indicated by shaded areas.

economic data and regional reports are analyzed. Members create economic projections and consider alternative scenarios. At the beginning of the actual meeting, the staff presents an overview of the current economic situation. Members then discuss policy options and vote on the proposed measures.[18] We use unrevised data available at the time of deliberations for our estimations to ensure that data revisions do not impact the results, focusing solely on the FOMC's real-time information set (Orphanides, 2001). We have previously documented that the literature finds central bank texts to provide additional informational value for understanding the monetary policy stance. While traditional Taylor rule determinants focus attention on how staff or market expectations, and past economic and inflation performance impact FOMC decisions, topics discussed and sentiment expressed in FOMC minutes, latent variables, likely also enter the mix when policy rate announcements are made.

First, we aim to analyze whether FOMC discussions on topics of the *Mandate* aggregate can explain monetary policy decisions, even after accounting for PCE inflation and the output gap. Second, we seek to understand how discussions of financial stability influence these decisions. We first estimate the following regression model based on a Taylor rule framework at the FOMC schedule frequency:

$$\begin{aligned} i_t = \alpha + \rho i_{t-1} + \beta^{SR}\mathbb{E}_t\big(\bar{\pi}_t^{1:4}\big) + \gamma^{SR}\mathbb{E}_t\big(\bar{y}_t^{1:4}\big) + \theta_{TP^M} TP_t^M + \theta_{TP^{FS}} TP_t^{FS} \\ + \theta_{Sent^M} Sentiment_t^M + \theta_{Sent^{FS}} Sentiment_t^{FS} + u_t \end{aligned} \tag{3}$$

As the dependent variable i_t, we use the effective Federal Funds Rate (FFR), and we model interest rate movements as influenced by an interest rate smoothing component and the variables of interest. To approximate the FOMC's real-time expectations on inflation and the output gap, we use the Greenbook forecasts available at each FOMC meeting. These forecasts are widely used in the literature as measures of the information available to the FOMC members when the policy decision is made (Coibion and Gorodnichenko, 2012). For each meeting t, we calculate the arithmetic mean of the Greenbook forecasts for the first four quarters following the quarter in which the meeting takes place. Specifically, $\mathbb{E}_t(\bar{\pi}_t^{1:4})$

[18]Since December 2004, detailed minutes have been published three weeks after the meeting (previously after the subsequent FOMC meeting), excluding any information revealed afterward (Danker and Luecke, 2005). This delay is irrelevant to our analysis because, unlike studies such as Tadle (2022), we do not focus on market reactions to the minutes but rather use the content to better understand whether, ex-post, FOMC deliberations influence the setting of the monetary policy stance.

is the average of the core PCE inflation forecasts for horizons $h = 1, \ldots, 4$, while $\mathbb{E}_t(\bar{y}_t^{1:4})$ is the analogous average of the output gap forecasts, as defined in equations (4) and (5).

$$\mathbb{E}_t\big(\bar{\pi}_t^{1:4}\big) = \frac{1}{4}\sum_{h=1}^{4}\mathbb{E}_t\Big(\pi_t^{q=h}\Big) \tag{4}$$

$$\mathbb{E}_t\big(\bar{y}_t^{1:4}\big) = \frac{1}{4}\sum_{h=1}^{4}\mathbb{E}_t\Big(y_t^{q=h}\Big) \tag{5}$$

Equation (3) is estimated in reduced form, so that β^{SR} and γ^{SR} represent the short-run responses to the inflation and output gap forecasts.[19]

TP_t^M and $Sentiment_t^M$ denote the topic proportions and the sentiment indicator for the *Mandate* aggregate, respectively. Analogously, TP_t^{FS} and $Sentiment_t^{FS}$ denote the corresponding variables for the *Financial Stability* aggregate. Table 3 reports descriptive statistics for these variables across the samples used in the empirical analysis. Panel A covers the full sample of FOMC meetings analyzed in Sections 3.2 and 3.3, whereas Panel B covers the Taylor rule estimation sample.[20] The distributions in Panels A and B are broadly similar, although financial stability discussions are somewhat more prevalent and their average sentiment is more negative in the Taylor rule estimation sample.

---

[19]Under the standard partial-adjustment interpretation of interest rate smoothing, the corresponding long-run coefficients are given by

$$\beta^{LR} = \frac{\beta^{SR}}{1-\rho} \qquad \text{and} \qquad \gamma^{LR} = \frac{\gamma^{SR}}{1-\rho}.$$

The long-run intercept is analogously obtained as $\alpha^{LR} = \frac{\alpha}{(1-\rho)}$.

[20]We restrict the Taylor-rule estimation sample to 2000–2018 and the press conference estimation sample to 2011–2018 because the required Greenbook forecasts were available to us for this period when we began working on the paper.

Table 3: Descriptive Statistics for Topic Proportions and Sentiment Indicators

Variable	N	Mean	Std. Dev.	Minimum	Maximum
Panel A: 1993–2022					
Topic Proportion *Mandate*	240	0.6383	0.1332	0.1667	0.9494
Topic Proportion *Financial Stability*	240	0.1254	0.1053	0.0000	0.7321
Sentiment *Mandate*	240	-0.0805	0.1952	-1.0000	0.4444
Sentiment *Financial Stability*	240	-0.0107	0.3553	-1.0000	1.0000
Panel B: Taylor Rule Sample (2000–2018)					
Topic Proportion *Mandate*	152	0.6106	0.1285	0.1667	0.8958
Topic Proportion *Financial Stability*	152	0.1503	0.1113	0.0000	0.7321
Sentiment *Mandate*	152	-0.0729	0.2141	-1.0000	0.4444
Sentiment *Financial Stability*	152	-0.0704	0.3348	-1.0000	1.0000
Panel C: Press Conference Sample (2011–2018)					
Topic Proportion *Mandate*	32	0.6533	0.1245	0.3263	0.8667
Topic Proportion *Financial Stability*	32	0.1476	0.0693	0.0381	0.4054
Sentiment *Mandate*	32	-0.0018	0.1648	-0.3810	0.3485
Sentiment *Financial Stability*	32	-0.0336	0.2686	-0.7000	0.4000

Note: The table reports the number of observations, mean, standard deviation, minimum, and maximum of the topic proportions and sentiment indicators used in the empirical analysis. Panel A covers our full time series of FOMC meetings from 1993 to 2022. Panel B covers the sample of 152 meetings used in the Taylor rule estimations. Panel C covers the sample of 32 meetings with Chair press conferences from April 2011 to December 2018. Topic proportions measure the share of FOMC discussions assigned to the respective topic aggregate. Higher sentiment values indicate more positive sentiment.

We do not interpret the estimated coefficients in terms of a full-unit change. Instead, we express the effects in terms of a 0.01-unit increase in the respective explanatory variable. For the topic proportions, this corresponds to a one-percentage-point increase in the share of FOMC discussions assigned to the respective topic aggregate. Following Oet and Lyytinen (2017), we expect θ_{TP^M} and $\theta_{TP^{FS}}$ to be negative, as greater discussion prevalence may reflect heightened concerns and a greater perceived need for monetary accommodation. By contrast, we expect θ_{Sent^M} and $\theta_{Sent^{FS}}$ to be positive, as more positive sentiment indicates fewer concerns when increasing interest rates.

A limitation when estimating Taylor rules to model monetary policy behavior is the ZLB, which compelled the Fed to implement unconventional policies to influence long-term interest

rates. Although shadow rates are often used for this period, it is important to recognize that these rates are estimates, not policy parameters directly controlled by the FOMC. We estimate a second specification to assess the impact of content and sentiment variables on monetary policy across these two different regimes. To capture the effect of the ZLB, we augment equation (3) by adding a ZLB dummy D_t^{ZLB}, equal to one during the ZLB period, with coefficient φ_{ZLB} to form interaction terms with the topic proportions and sentiment variables.[21] This yields the following equation (6):

$$\begin{aligned} i_t = \alpha + \rho i_{t-1} + \beta^{SR}\mathbb{E}_t\big(\bar{\pi}_t^{1:4}\big) + \gamma^{SR}\mathbb{E}_t\big(\bar{y}_t^{1:4}\big) + \Big[\theta_{TP^M} TP_t^M + \theta_{TP^{FS}} TP_t^{FS} + \\ \theta_{Sent^M} Sentiment_t^M + \theta_{Sent^{FS}} Sentiment_t^{FS}\Big] + \varphi_{ZLB} D_t^{ZLB} + D_t^{ZLB} * \Big[\varphi_{TP^M} TP_t^M + \\ \varphi_{TP^{FS}} TP_t^{FS} + \varphi_{Sent^M} Sentiment_t^M + \varphi_{Sent^{FS}} Sentiment_t^{FS}\Big] + u_t \end{aligned} \tag{6}$$

Here, φ_{ZLB} denotes the coefficient on the ZLB dummy, while φ_{TP^M}, $\varphi_{TP^{FS}}$, φ_{Sent^M}, and $\varphi_{Sent^{FS}}$ denote the coefficients on the respective interactions with the ZLB dummy. Consequently, the total effect of each topic proportion or sentiment indicator during the ZLB period is given by the sum of its baseline coefficient and the corresponding interaction coefficient.

Finally, we analyze how the topic proportions and sentiment of discussions related to the *Mandate* and *Financial Stability* aggregates influence the Fed Chair's monetary policy communication during the press conferences following FOMC meetings. To this end, we estimate sentiment regressions, that is, we use sentiment variables on the LHS and RHS of the regression specification (Kanelis and Siklos, 2025). We use RoBERTa to measure sentiment from the introductory statement and the Q&A session. For the latter, we restrict the analysis to the Chair's answers.

$$SR_t^{Statement} = \frac{\#Positive_t^{Sentence} - \#Negative_t^{Sentence}}{\#Positive_t^{Sentence} + \#Neutral_t^{Sentence} + \#Negative_t^{Sentence}} \tag{7}$$

$$SR_t^{Answers} = \frac{\#Positive_t^{Answers} - \#Negative_t^{Answers}}{\#Positive_t^{Answers} + \#Neutral_t^{Answers} + \#Negative_t^{Answers}} \tag{8}$$

[21] The ZLB period is from December 2008 until December 2015.

We calculate the sentiment of the introductory statement (equation (7)) by subtracting the number of negative sentences from the number of positive sentences and dividing the result by the total number of sentences, that is, the sum of positive, neutral, and negative sentences. We analogously calculate the sentiment of the Q&A session (equation (8)), but classify each answer as a whole rather than individual sentences. These sentiment measures are then used as dependent variables in the sentiment regression (SR_t) specifications.

$$\begin{aligned} SR_t^D = \alpha + \beta \mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right) + \gamma \mathbb{E}_t\left(\bar{y}_t^{1:4}\right) + \theta_{TP^M} TP_t^M + \theta_{TP^{FS}} TP_t^{FS} + \\ \theta_{Sent^M} Sentiment_t^M + \theta_{Sent^{FS}} Sentiment_t^{FS} + \zeta i_{t-1} + u_t \end{aligned} \tag{9}$$

where $D \in \{Statement, Answers\}$. As in equations (3) and (6), we use Greenbook forecasts of core PCE inflation and the output gap, together with the topic proportions and sentiment indicators derived from the discussions in the minutes on topics from the *Mandate* and *Financial Stability* aggregates.[22] Panel C of Table 3 shows that the topic proportions in the press conference sample are broadly similar to those in the Taylor rule estimation sample reported in Panel B, although mandate-related discussions are somewhat more prevalent and the average sentiment indicators are closer to zero. We additionally include the lagged interest rate to control for the previous monetary policy stance.[23] We expect the same signs as before, given that monetary policy communication aims to influence market interest rates in a manner similar to interest rate policies.

5. Results

5.1. Monetary Policy Action

We first examine monetary policy actions as reflected in interest rate decisions over the period from 2000 to 2018. The sample comprises 152 FOMC meetings, and we estimate the Taylor rule specifications using ordinary least squares.[24] Estimates are reported in Table 4.

[22]We do not include ZLB interaction terms because the press conferences started only in 2011. Consequently, the sample contains no pre-ZLB observations and only a relatively short post-ZLB period.

[23]We use the lagged rather than the contemporaneous interest rate because the current policy rate decision and the associated communication are jointly determined at the same FOMC meeting. Conditioning on i_t could therefore introduce simultaneity and absorb part of the estimated effects of the topic and sentiment indicators.

[24]The topic proportions and sentiment indicators included in the regressions are measured rather than directly observed and therefore constitute generated regressors. In the regressions, we treat these measured

In column (1), we present the baseline Taylor rule and augment it stepwise with the topic proportion variables (column (2)) and the sentiment variables (column (3)) for the topic aggregates *Mandate* and *Financial Stability*, respectively. For all model specifications, we estimate a high interest rate smoothing parameter (ρ). The validity of the Taylor principle is confirmed by the steady-state estimates in the upper part of the table, consistent with the literature (Carvalho et al., 2021). Furthermore, in the short-run, the output gap also affects the FFR in most specifications in a direction consistent with theory.

Now, we consider the textual variables derived from the minutes to test our hypotheses. In specification (2), we find statistically significant negative effects of the topic proportions pertaining to *Mandate* (θ_{TP^M}) on the FFR. This is particularly interesting because the variable captures relative discussion shares. An increase in this variable translates into a declining discussion proportion of all other topics during the FOMC meeting. Specifically, on average, a one-percentage-point increase in TP^M is associated with a short-run decrease in the FFR of approximately 0.76 basis points. When the FOMC places a stronger focus on mandate topics, it becomes more likely that the Fed will reduce interest rates.

The results become more mixed after augmenting the model with the sentiment indicators (specification (3)). On average, a more positive discussion of the *Mandate* topic (θ_{SentM}) generates a higher FFR. To illustrate the combined marginal effects, a one-percentage-point decline in the mandate topic proportion together with a 0.01-unit increase in mandate sentiment is associated with a short-run increase in the FFR of approximately 1.09 basis points. At the same time, we estimate no significant effect of the content and sentiment variables for *Financial Stability* ($\theta_{Sent^{FS}}$), suggesting that financial stability is less relevant for interest rate policy in this context. However, we must consider that a significant part of financial stability discussions occurred following the GFC when monetary policymakers in the US were also confronted with the ZLB.

This raises the question of how the GFC influenced the FOMC's behavior. To account for the ZLB, we estimate equation (6) and present the results in column (4) of table (4). To ensure that we account for unconventional monetary policy instruments, we replace the dependent variable with the shadow rate of Wu and Xia (2016) and present the results in column (5).

---

variables as fixed. The resulting inference is therefore conditional on the measured topic proportion and sentiment variables and may understate total sampling uncertainty. We report bootstrap standard errors based on 1,000 replications to account for sampling uncertainty in the regressions. However, the uncertainty associated with measuring the topic proportions and sentiment indicators is not taken into account.

Table 4: Taylor Rule Estimates (2000–2018)

	FFR				Shadow Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Long-run estimates						
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	1.320	1.875*	3.039*	5.619**	6.882	6.838
	(1.378)	(1.027)	(1.779)	(2.410)	(4.976)	(6.182)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.653*	0.582**	0.465	0.849	2.560	3.053
	(0.343)	(0.254)	(0.374)	(0.646)	(2.440)	(3.533)
Constant	0.441	7.401*	6.748	6.686	17.900	27.277
	(2.635)	(3.935)	(6.098)	(6.860)	(17.661)	(31.021)
Short-run estimates						
i_{t-1}	0.947***	0.929***	0.949***	0.945***	0.970***	0.976***
	(0.022)	(0.025)	(0.022)	(0.027)	(0.024)	(0.025)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.069	0.133	0.156	0.309***	0.207*	0.164
	(0.085)	(0.095)	(0.109)	(0.118)	(0.111)	(0.112)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.034***	0.041***	0.024	0.047*	0.077***	0.073***
	(0.012)	(0.014)	(0.018)	(0.025)	(0.025)	(0.026)
TP_t^M		-0.755**	-0.651**	-0.914**	-1.015**	-0.991**
		(0.300)	(0.325)	(0.448)	(0.423)	(0.427)
TP_t^{FS}		-0.682*	-0.367	-1.358**	-1.260**	-1.171**
		(0.368)	(0.344)	(0.547)	(0.510)	(0.511)
$Sentiment_t^M$			0.438**	0.713***	0.681***	0.636***
			(0.171)	(0.189)	(0.191)	(0.176)
$Sentiment_t^{FS}$			-0.026	-0.033	0.002	0.004
			(0.102)	(0.117)	(0.111)	(0.110)
VIX_t						-0.004
						(0.007)
D_t^{ZLB}				-0.630*	-1.078***	-1.100***
				(0.334)	(0.395)	(0.397)
$D_t^{ZLB} \times TP_t^M$				0.786*	1.464***	1.485***
				(0.444)	(0.497)	(0.497)
$D_t^{ZLB} \times TP_t^{FS}$				1.783**	2.605***	2.617***
				(0.719)	(0.820)	(0.810)
$D_t^{ZLB} \times Sentiment_t^M$				-0.963***	-0.850***	-0.835***
				(0.261)	(0.288)	(0.278)
$D_t^{ZLB} \times Sentiment_t^{FS}$				0.060	0.179	0.152
				(0.158)	(0.181)	(0.192)
Constant	0.023	0.526**	0.347	0.368	0.538	0.652*
	(0.134)	(0.249)	(0.308)	(0.383)	(0.376)	(0.396)
Wald Tests (χ^2)						
$\theta_{TP^M} + \varphi_{TP^M} = 0$				0.429	1.798	2.167
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$				1.597	6.290**	6.810***
$\theta_{Sent^M} + \varphi_{Sent^M} = 0$				1.586	0.611	0.811
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$				0.059	1.578	1.056
Sample	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018
N	152	152	152	152	152	152
$Adj.R^2$	0.971	0.972	0.973	0.977	0.984	0.984

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018. Columns (1)–(4) use the effective federal funds rate as the dependent variable. Columns (5)–(6) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period, while column (6) additionally controls for the VIX. $\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate and financial stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment measures, constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. Long-run estimates are obtained by dividing the corresponding short-run coefficients by $1-\rho$, where ρ denotes the estimated coefficient on the lagged dependent variable. The Wald test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Controlling for the ZLB period confirms our previous findings: outside the ZLB, increases in the discussion shares around *Mandate* and *Financial Stability* lead, on average, to interest rate reductions found elsewhere in the literature (Istrefi et al., 2023). Nevertheless, the ZLB period (φ_{ZLB}) does exert some influence.

To evaluate the impact of textual variables during the ZLB period, we use Wald tests to analyze the effect of variables conditional on $D_t^{ZLB} = 1$. Looking at column (4), the topic proportions and sentiment variables for *Mandate* and *Financial Stability* lose their significance in explaining interest rate movements. This aligns with our priors, as interest rates were downwardly rigid during the ZLB. Simultaneously, the FOMC was hesitant to raise interest rates prematurely despite post-GFC economic improvements indicated by positive sentiment in the minutes (see Figure (3)). Consequently, it is unsurprising that textual variables fail to explain changes in the FFR, as the FOMC intentionally kept the rate unchanged.

In contrast, when analyzing unconventional monetary policy instruments proxied by the shadow rate, we find that an increase in the proportion of financial stability topics discussed by the FOMC correlates with tighter financing conditions ($\theta_{TP^{FS}} + \varphi_{TP^{FS}}$). This relationship likely reflects concerns that prolonged periods of low interest rates and unconventional policy tools may undermine financial stability (Kashyap and Stein, 2023). As Yellen (2014) notes, persistently low rates can incentivize "reach for yield" behavior, leading investors to take on excessive risk and potentially destabilizing the financial system.[25]

Importantly, we provide novel evidence that greater attention to financial stability during FOMC deliberations is associated with monetary policy tightening at the ZLB, even after controlling for sentiment. To our knowledge, this relationship has not been documented previously in the literature. This finding extends the evidence in Istrefi et al. (2023), who document an accommodative relationship between financial stability communication and monetary policy before the GFC but find no corresponding policy signal in the post-2007 period. To ensure that our financial stability indicators derived from FOMC deliberations provide informational value beyond publicly available measures, we re-estimate equation (6) with the VIX as an additional control in column (6). The estimated results remain robust, indicating that the textual variables from the minutes capture a unique perspective of the FOMC.

[25]Similar arguments are presented by Tarullo (2014).

5.2. Monetary Policy Communication

After analyzing the monetary policy actions of the FOMC, we now examine whether immediate monetary policy communication is affected. Since April 2011, the FOMC has held press conferences, initially four times a year, and eventually, starting in January 2019, after every FOMC meeting. In these press conferences, the FOMC Chair addresses the media and international markets after the conclusion of the meeting and reads a prepared statement. In this statement, the Fed Chair explains the monetary policy decisions and provides an overview of the current economic situation and the outlook. Afterward, journalists have the opportunity to ask questions.

We examine whether the FOMC's discussions of *Mandate* and *Financial Stability* are associated with the sentiment of the Chair's introductory statement. We estimate equation (9) using the sentiment measure for the Chair's introductory statement defined in equation (7) and report the results in Table 5. The sample includes all post-meeting press conferences through December 2018.

Looking at column (1), we find a statistically significant effect only for the expected output gap variable. A higher positive output gap or a smaller negative one generates increased optimism in the statement. Once again, rather than estimating these figures, we rely on the official values from the Greenbook forecasts to ensure that only information available at the time is used. At first glance, the lack of statistical significance for the expected core PCE inflation rate is surprising. Still, it is understandable in light of the observed period during which the FOMC was particularly focused on elevated unemployment and substantial labor market slack in the aftermath of the GFC. This is also consistent with Bohl et al. (2023), who similarly find that expectations regarding unemployment developments after the GFC influenced the tone of speeches.

Incorporating the text-based variables derived from the minutes yields several notable results. As columns (2) and (3) show, the Greenbook forecasts of inflation and the output gap are not statistically significant once the topic proportions and sentiment indicators are included. By contrast, the sentiment of mandate-related discussions (θ_{Sent^M}) is positively and statistically significantly associated with the sentiment of the Chair's introductory statement. In column (3), a one-percentage-point increase in net sentiment during mandate-related discussions is associated with an approximately 0.24-percentage-point increase in the sentiment of the Chair's introductory statement. This result makes sense, as the discussions during the FOMC meeting are much more in-depth than just setting interest rates

Table 5: Sentiment Regression on Fed Chair's Statement and Q&A Session Responses (2011-2018)

	Statement			Q&A Session		
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.111	0.115	0.141	0.024	0.064	-0.043
	(0.113)	(0.120)	(0.140)	(0.052)	(0.070)	(0.075)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.018**	0.003	0.007	0.006	0.002	-0.012
	(0.008)	(0.014)	(0.015)	(0.005)	(0.006)	(0.009)
TP_t^M		0.088	0.073		0.093	0.155
		(0.193)	(0.198)		(0.151)	(0.152)
TP_t^{FS}		-0.275	-0.239		-0.050	-0.198
		(0.241)	(0.263)		(0.183)	(0.195)
$Sentiment_t^M$		0.265**	0.243**		-0.028	0.062
		(0.114)	(0.116)		(0.078)	(0.089)
$Sentiment_t^{FS}$		-0.041	-0.034		0.075	0.048
		(0.054)	(0.059)		(0.046)	(0.043)
i_{t-1}			-0.008			0.033**
			(0.016)			(0.013)
$Constant$	-0.062	-0.104	-0.142	-0.059	-0.179	-0.022
	(0.194)	(0.278)	(0.299)	(0.087)	(0.157)	(0.140)
Sample	2011-2018	2011-2018	2011-2018	2011-2018	2011-2018	2011-2018
N	32	32	32	32	32	32
$Adj.R^2$	0.344	0.423	0.404	0.040	0.063	0.245

Note: The table reports OLS estimates of equation (9) for the period from April 2011 to December 2018. Each observation corresponds to one press conference. The dependent variable is the sentiment of the Chair's introductory statement in columns (1)–(3) and the sentiment of the Chair's responses during the Q&A session in columns (4)–(6). $\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. TP_t^M and TP_t^{FS} are the proportions of FOMC discussions devoted to mandate objectives and financial stability, while $Sentiment_t^M$ and $Sentiment_t^{FS}$ are the corresponding topic-specific sentiment measures. i_{t-1} denotes the policy rate at the previous FOMC meeting, measured by the effective federal funds rate and replaced by the shadow rate of Wu and Xia (2016) during the ZLB period. Sentiment is measured using the fine-tuned RoBERTa model described in Section 3.3. Bootstrapped standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

based on expected changes in specific macroeconomic variables.

Our analysis of monetary policy decisions and communication reveals an interesting feature of financial stability discussions. While an increase in such discussions is clearly associated with monetary policy behavior, we find no statistically significant association between financial stability discussions in the minutes and the sentiment of subsequent monetary policy communication. This asymmetry reflects a dilemma that may be explained by the Fed's mandate: although financial stability is not formally an element of the mandate, it nevertheless affects monetary policy through various channels (Kashyap and Stein, 2023).

Finally, we examine the effect of the content of discussions during FOMC meetings on Q&A sessions during press conferences. We once again estimate equation (9) with the sentiment indicator from equation (8) as the dependent variable and present the results in columns (4) - (6) of Table (5). We find no systematic relationship between the expected and discussed macroeconomic developments during the FOMC meeting and the sentiment of the responses. However, column (6) shows a positive and statistically significant association between the lagged policy rate and the sentiment of the answers. One possible explanation for the otherwise weak relationship is that, unlike the statement, the answers are less scripted and respond directly to journalists' questions.

5.3. Anecdotal Evidence from the Transcripts

To better understand our results and the reversal in the estimated total effect of TP^{FS} during the ZLB period, we apply a narrative approach in the spirit of Romer and Romer (2023). We read the FOMC transcripts from 2007 through the first rate hike after the GFC to examine how these discussions evolved. Published with a five-year lag, these transcripts provide a nearly verbatim account of the Committee's deliberations. Table (6) presents selected examples of how FOMC participants discussed financial stability considerations during policy deliberations, with additional examples reported in Appendix C. Although financial stability is not formally part of the Federal Reserve's dual mandate, the excerpts show that it entered the FOMC's deliberations over the appropriate monetary policy stance, as illustrated by examples (4) and (5) of Table (6).

In 2007, the FOMC adopted a more accommodative policy stance in response to weaknesses in the housing sector and strain in financial markets, which created downside risks to economic activity (Table (C1), example (1)). This response was consistent with the prevailing view that reductions in the FFR represented a conventional monetary policy response to financial crises, as Bernanke noted during the discussions (Table (6), example (1)). However,

Table 6: Illustrative Quotes from the FOMC Transcripts on Financial Stability

Excerpt	Date	FOMC Member	Quote
1	October 30-31, 2007	Ben Bernanke	"For what it is worth, 75 basis points of easing has been pretty much the standard Fed medicine for financial crises ever since 1970 or so; in that respect we are in good company."
2	March 16, 2010	Donald L. Kohn	"I don't see, so far, low rates as having greatly adverse consequences for financial stability, though I agree we need to monitor it. The bond spreads, equity prices, I think, are still in line with fundamentals. And given my expectation that both inflation and output and employment will remain below our objectives for a long period of time, I would have a very strong preference for using supervision and regulation to counter any hints of instability in financial markets or dangerous situations in financial markets rather than have those measures fall even further below our objectives."
3	March 16, 2010	Thomas M. Hoenig	"The reason to begin tightening policy soon is not simply or even primarily to prevent the issues around inflationary pressures immediately. That's longer term and important. Rather it is to prevent future buildup of financial imbalances and reduce the risks to longer-run macroeconomic and financial stability that come from maintaining a commitment to exceptionally low rates for an extended period of time."
4	October 28-29, 2014	Stanley Fischer	"On financial stability, that's a very tough one indeed. The current constitutional document is based in the legal system, in laws relating to our mandate. I don't think financial stability is equally based. There's a slight problem there. Nonetheless, I liked a lot of what I heard around the table. I think that the key issue here is not how the Fed deals with financial stability but how the FOMC deals with financial stability. And that means we control the interest rate, and do we ever use the interest rate to deal with financial-stability issues? Now, President Evans used the Greenspan line: "Well, if you were just going to use the interest rate when the situation becomes untenable, you'd have to do so much to raise the interest rate that you'd destroy the economy." That prejudged the answer. We can't do it. We wait for the bubble to pop, and then we mop it up. The issue is, do we take financial-stability considerations into account in our interest rate decisions? We sometimes do, and then Lael pointed out that sometimes we do it right at the beginning and not right at the end. I think we ought to work a little bit more on that, but we cannot develop the statement that implies it's on the same level as inflation and unemployment."
5	December 15-16, 2015	Stanley Fischer	"Why move now? Why not wait longer? First, as the Chair has emphasized, our actions become effective with a lag. Second, there are some signs of accumulating financial stability problems. And, third, the signal we will be sending will reinforce the fact that our economic situation is continuing to normalize. [. . .] So, to conclude, we're setting out on a new journey with the same goals we've always had—maximum employment, stable prices, and financial stability."

Note: The table reports selected excerpts from the FOMC transcripts illustrating how Committee members discussed financial stability in the context of monetary policy decisions. The excerpts were identified using a manual narrative approach inspired by Romer and Romer (2023) and provide qualitative evidence consistent with our econometric results. Additional excerpts are reported in Appendix C.

this perspective shifted during the ZLB period, as the prolonged accommodative policy itself became a source of concern for financial stability.

At the beginning of the ZLB period, FOMC participants differed in their assessments of these risks. Kohn saw little evidence that low interest rates had generated significant adverse consequences for financial stability and favored addressing emerging risks through supervision and regulation. By contrast, Hoenig argued that an earlier tightening of monetary policy could help prevent the buildup of financial imbalances (Table (6), examples (2) and (3)). As the ZLB period progressed, concerns about financial stability risks arising from persistently accommodative monetary policy became more pronounced. Participants warned that large-scale asset purchases could encourage excessive risk-taking (Table (C1), example (3)) and that prolonged low interest rates and forward guidance could foster reaching for yield, asset mispricing, and deteriorating underwriting standards (Table (C1), examples (4) and (6)–(8)). In example (4) of Table (6), Stanley Fischer emphasized that the FOMC sometimes considered financial stability in its interest rate decisions, while also noting that it could not be placed on the same footing as the statutory dual mandate. This distinction underscores the policy relevance of financial stability without treating it as an equivalent objective.

Later in the ZLB period, the debate increasingly focused on the appropriate timing of policy normalization. Participants recognized financial stability risks on both sides: prolonged accommodation could encourage excessive risk-taking, while tightening policy too early or too abruptly could itself destabilize financial markets and the broader economy. These competing considerations are reflected in the discussion of policy exit strategies reported in Table 6, example (5) and Table (C1), examples (10) – (12).

5.4. Analysis with Explainable AI

The use of sophisticated models is accompanied by a loss of explainability and interpretability of the inner workings of the model. This critique of deep learning models is widespread. In this study, we address it by introducing, as a further methodological contribution, an xAI approach that makes our sentiment analysis transparent. Using this approach, we derive a new dictionary that replicates our results and thus provides explainability. The dictionary provides a more transparent representation of our LLM classification approach and serves as a robustness check for our analysis.

We construct the dictionary from SHAP values so that it estimates sentiment similarly to our fine-tuned RoBERTa model. Using a word-level tokenizer, we assign each word a single SHAP value, avoiding aggregation across subword tokens and simplifying interpretation.

We lemmatize the corpus and compute, for each lemma, the average SHAP value across all occurrences. To ensure reliability, we exclude words appearing fewer than ten times and remove stopwords. Consequently, each dictionary entry has average SHAP values for the positive, negative, and neutral sentiment categories.

We then assign each word to the sentiment category with the highest average SHAP value. Words with higher average positive (negative) SHAP values are classified as positive (negative). From these two lists, we retain only the top 20% by average contribution, ensuring that each category contains only the highest-contributing words. This procedure improves the dictionary's accuracy in estimating sentiment while underscoring the limitations of static dictionaries compared to the context-aware capabilities of language models. Using the held-out sample, we find that our SHAP-derived dictionary achieves modestly higher accuracy and overall classification performance than the hand-crafted dictionary by Loughran and McDonald (2011) in the context of FOMC deliberations.

Table (7) compares the estimates obtained using our constructed dictionary in columns (1) and (2) with the baseline estimates from the fine-tuned RoBERTa model in columns (3) and (4). Although the coefficient magnitudes and significance levels differ to some extent, the main qualitative conclusions remain unchanged. In particular, mandate sentiment remains positively associated with the policy rate outside the ZLB. Moreover, the total effect of financial stability topic proportions on the shadow rate during the ZLB remains positive and statistically significant. Appendix E provides further details on the construction of the dictionary, presents the complete word list, and reports its evaluation results.

5.5. Robustness Checks and Extensions

In this section, we conduct a series of robustness checks and extensions that modify key measurement choices, model specifications, and empirical assumptions. Combining machine learning based measurement with Taylor rule estimation gives researchers discretion in choosing parameters and design specifications. We therefore examine whether our results depend on these choices. We organize the robustness analysis in the same sequence as the main analysis, moving from topic modeling and sentiment measurement to econometric specifications. Appendix B presents the corresponding tables and figures.

First, we examine whether our findings depend on the aggregation of STM topics into broader aggregates *Mandate* and *Financial Stability*. We implement two alternative aggregation schemes.

Table 7: Taylor Rule Estimates with SHAP Dictionary (2000–2018)

	SHAP Dictionary		RoBERTa Model	
	FFR	Shadow Rate	FFR	Shadow Rate
	(1)	(2)	(3)	(4)
Short-run estimates				
i_{t-1}	0.935***	0.968***	0.945***	0.970***
	(0.027)	(0.023)	(0.027)	(0.024)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.369***	0.245**	0.309***	0.207*
	(0.134)	(0.122)	(0.118)	(0.111)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.059**	0.090***	0.047*	0.077***
	(0.025)	(0.026)	(0.025)	(0.025)
TP_t^M	-1.024**	-1.101***	-0.914**	-1.015**
	(0.430)	(0.410)	(0.448)	(0.423)
TP_t^{FS}	-1.319**	-1.140*	-1.358**	-1.260**
	(0.649)	(0.600)	(0.547)	(0.510)
$Sentiment_t^M$	0.509**	0.551**	0.713***	0.681***
	(0.241)	(0.243)	(0.189)	(0.191)
$Sentiment_t^{FS}$	0.106	0.131	-0.033	0.002
	(0.138)	(0.136)	(0.117)	(0.111)
D_t^{ZLB}	-0.474	-0.872**	-0.630*	-1.078***
	(0.345)	(0.384)	(0.334)	(0.395)
$D_t^{ZLB} \times TP_t^M$	0.807*	1.455***	0.786*	1.464***
	(0.445)	(0.499)	(0.444)	(0.497)
$D_t^{ZLB} \times TP_t^{FS}$	1.914**	2.489***	1.783**	2.605***
	(0.857)	(0.888)	(0.719)	(0.820)
$D_t^{ZLB} \times Sentiment_t^M$	-0.705**	-0.812**	-0.963***	-0.850***
	(0.326)	(0.356)	(0.261)	(0.288)
$D_t^{ZLB} \times Sentiment_t^{FS}$	-0.144	0.075	0.060	0.179
	(0.215)	(0.274)	(0.158)	(0.181)
Constant	0.253	0.409	0.368	0.538
	(0.418)	(0.406)	(0.383)	(0.376)
Wald Tests (χ^2)				
$\theta_{TP^M} + \varphi_{TP^M} = 0$	0.734	0.999	0.429	1.798
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$	1.955	6.292**	1.597	6.290**
$\theta_{Sent^M} + \varphi_{Sent^M} = 0$	0.688	0.927	1.586	0.611
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$	0.049	0.732	0.059	1.578
Sample	2000–2018	2000–2018	2000–2018	2000–2018
N	152	152	152	152
$Adj.R^2$	0.975	0.983	0.977	0.984

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018. Columns (1) and (3) use the effective federal funds rate as the dependent variable, whereas columns (2) and (4) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period. $\mathbb{E}_t(\bar{\pi}_t^{1:4})$ and $\mathbb{E}_t(\bar{y}_t^{1:4})$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate-related and financial stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment indicators. In columns (1)–(2), these indicators are constructed using the SHAP-derived dictionary developed from the fine-tuned RoBERTa model. Columns (3)–(4) report the corresponding baseline estimates using the sentence-level sentiment classifications generated directly by the fine-tuned RoBERTa model described in Section 3.3. Appendix E provides further details on the construction of the dictionary, presents the resulting word lists, and reports the additional evaluation results. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. The Wald test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Under the first scheme, we construct narrower versions of the baseline categories. For *Mandate*, we restrict the aggregation to *Labor Market Conditions*, *Dual-Mandate Objectives*, *Inflation Expectations*, and *Inflation Outlook*. For *Financial Stability*, we exclude *Housing and Mortgages* when calculating topic proportions and sentiment. This restriction follows the logic of Istrefi et al. (2023), who treat housing separately from other financial topics because housing market discussions may also capture broader macroeconomic developments in addition to financial stability concerns. The results remain robust under these narrower definitions. Appendix B.1 reports the topic aggregation using the narrow definition and the corresponding estimation results.

Under the second scheme, we replace the aggregate *Mandate* category with separate measures of *Economic Outlook* and *Monetary Policy Inclination*, following the distinction used by Ehrmann and Fratzscher (2007). *Economic Outlook* captures discussions on real economic activity, demand, growth, and macroeconomic risks. *Monetary Policy Inclination* captures discussions on inflation and assessments and discussions on the monetary policy stance. Using this approach, we also incorporate discussions on monetary policy that are not part of the *Mandate* aggregate. This decomposition also connects to recent work that applies these concepts, like Wischnewsky et al. (2021). Our findings remain robust. Outside the ZLB, a higher prevalence of financial stability discussions continues to be associated with a lower policy rate, whereas during the ZLB period it is associated with a higher shadow rate. The alternative decomposition additionally shows that more positive sentiment in both *Economic Outlook* and *Monetary Policy Inclination* is associated with higher policy rates outside the ZLB. Appendix B.1 reports the topic aggregation using this alternative decomposition and the corresponding estimation results.

Second, we vary the most central parameter of the STM: K, the number of latent topics identified by the model. To assess whether the results depend on the baseline choice of $K = 21$, we estimate alternative Pareto-efficient topic specifications along the semantic coherence–exclusivity frontier. Specifically, we estimate the STM for $K \in \{18, 22, 25, 26, 27, 29\}$ and aggregate the resulting topics using the same economic interpretation based on keywords and representative paragraphs as in the baseline specification. The results remain robust for $K \in \{18, 22, 25, 26, 27, 29\}$. Appendix B.2 illustrates the sensitivity of the estimates using coefficient plots.

Third, we relax the symmetry imposed by our net sentiment measure in the Taylor rule estimation, which assigns positive and negative sentences equal weights. Although this indi-

cator is common in the literature on central bank communication (Apel et al., 2022; Kanelis and Siklos, 2025), several studies from financial and monetary economics emphasize that negative information may carry greater informational content (Tetlock, 2007; Wischnewsky et al., 2021; Istrefi et al., 2023). We therefore first replace our sentiment indicator from Equation (2) with a negativity indicator and then estimate a more flexible specification that includes separate positivity and negativity indicators for the *Mandate* and *Financial Stability* topic aggregates. Wald tests assess whether the positive and negative coefficients are equal in magnitude and opposite in sign, both outside and during the ZLB period. The negativity-only specification confirms that more negative mandate-related sentiment is associated with a lower federal funds rate outside the ZLB, whereas negative financial stability sentiment has no significant additional effect. In the joint specification, however, Wald tests do not reject the symmetry assumption underlying the baseline sentiment measure. Hence, our main conclusions remain unchanged. The construction of the indicators, the full estimation results, and the Wald tests are reported in Appendix B.3.

Fourth, we allow the effect of sentiment to vary with the salience of the corresponding topic in the Taylor rule estimation. While the baseline specifications include topic proportions and sentiment indicators additively, we extend equation (6) by adding interactions between topic proportions and sentiment indicators, as well as their interaction with the ZLB indicator. This allows the marginal effect of sentiment to vary with topic prevalence and across policy regimes. Appendix B.4 reports the extended specification and results. The main findings remain robust. The interaction estimates further suggest that, conditional on financial stability topic prevalence, a more negative tone in financial stability discussions is associated with lower policy rates outside the ZLB. This supports our interpretation that topic prevalence captures the presence of concerns, while tone contains additional information about their intensity.

Fifth, we examine the sensitivity of our results to alternative empirical and econometric assumptions. Appendix B.5 reports several alternative specifications of equation (6): we replace the Greenbook output gap forecasts with Greenbook forecasts for the unemployment gap, vary the forecast horizon from four to two quarters, substitute the VIX with alternative financial indicators (NFCI, FSI, and KCFSI), augment the Taylor rule by adding a second lag of the dependent variable, and replace the shadow rate estimates of Wu and Xia (2016) with those of Krippner (2013). The results remain robust across all alternative specifications.

Finally, instead of modeling the regime shift through interaction terms with a ZLB indicator,

we estimate the Taylor rule separately for the non-ZLB and ZLB periods, using December 2008 as the breakpoint. As an additional robustness check, we restrict the ZLB subsample to January 2010–December 2015, excluding December 2008 and the year 2009to ensure that the results are not driven by the acute phase of the GFC and the associated severe U.S. recession. Our results remain robust across these subsample specifications (see Table B16).

6. Conclusions

Our findings show that the information contained in FOMC deliberations, as summarized in the minutes, helps explain monetary policy decisions and subsequent policy communication beyond conventional macroeconomic forecasts. A greater prevalence of discussions related to the *Mandate* aggregate is associated with a lower FFR. The sentiment of these discussions provides additional insights. The Taylor rule and sentiment regression estimates indicate that more positive sentiment in mandate-related discussions is associated with a higher policy rate and a more positive tone in the Chair's introductory statement following the meeting. Outside the ZLB, a greater prevalence of financial stability discussions is also associated with a lower policy rate.

Distinguishing between conventional and unconventional monetary policy substantially changes the interpretation of the financial stability results. Using the shadow rate during the ZLB period, we find that the total effect of financial stability topic proportions is positive. This finding is consistent with concerns among FOMC participants that prolonged monetary accommodation could encourage excessive risk-taking by financial market participants, potentially sowing the seeds of future financial instability-induced recessions. The narrative evidence from the FOMC transcripts supports this interpretation. By contrast, we find no statistically significant association between financial stability discussions in the minutes and the sentiment of either the Chair's introductory statement or the Chair's responses during the Q&A session. This asymmetry suggests that financial stability considerations enter monetary policy decisions without being communicated with equal prominence. Consequently, policy decisions cannot be fully understood from public communication alone.

Methodologically, our study contributes by constructing a novel dataset of annotated sentences from FOMC minutes and using it to fine-tune and evaluate RoBERTa for sentiment analysis. Extensive evaluation demonstrates that our approach outperforms established hand-crafted dictionary-based approaches. We also introduce SHAP values to the literature on central bank deliberations to make the model's classification more transparent and derive

a SHAP-based sentiment dictionary. The newly-created dictionary provides an interpretable robustness check and yields Taylor rule estimates that preserve the main qualitative conclusions obtained with the fine-tuned RoBERTa model. Our approach therefore combines the context-sensitive capabilities of language models with the transparency of a dictionary-based sentiment measure.

Our results contribute to the expanding literature on the role of financial stability considerations in shaping how the stance of monetary policy is set. Future research should harness modern NLP techniques to more precisely capture how various topics within FOMC discussions influence monetary policy decisions. A more granular understanding of these topic-specific effects could provide deeper insights into how financial stability and other considerations inform the FOMC's policy-setting behavior. Future studies can use complementary methodologies to deepen our understanding of how market participants process this additional information beyond what is explicitly communicated. A promising approach involves analyzing and comparing market reactions to statements and minutes, providing insights into how investors interpret the asymmetry between monetary policy decisions and their communication. Additionally, comparing central banks, such as the Eurosystem or the Bank of England, could also shed light on whether similar communication-policy asymmetries exist in different institutional frameworks. After all, the Fed's mandate, described as "muddled" by Orphanides (2024), gives it considerable latitude in determining the importance placed on the price stability objective. This differs from the mandates of the Eurosystem and the Bank of England. Hence, our findings may also have implications for how central bank mandates are designed.

References

Acosta, M., 2023. *A New Measure of Central Bank Transparency and Implications for the Effectiveness of Monetary Policy*. International Journal of Central Banking 19, 49–97.

Adrian, T., Liang, N., 2018. *Monetary Policy, Financial Conditions, and Financial Stability*. International Journal of Central Banking , 73–131.

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J.M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., Herrera, F., 2023. *Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence*. Information Fusion 99, 101805.

Apel, M., Blix Grimaldi, M., Hull, I., 2022. *How Much Information Do Monetary Policy Committees Disclose? Evidence from the FOMC's Minutes and Transcripts*. Journal of Money, Credit and Banking 54, 1459–1490.

Aruoba, S.B., Drechsel, T., 2026. *Identifying Monetary Policy Shocks: A Natural Language Approach*. American Economic Journal: Macroeconomics. Forthcoming.

Bauer, M.D., Bernanke, B.S., Milstein, E., 2023. *Risk Appetite and the Risk-Taking Channel of Monetary Policy*. Journal of Economic Perspectives 37, 77–100.

Beckmann, L., Debener, J., Kriebel, J., 2023. *Understanding the determinants of bond excess returns using explainable AI*. Journal of Business Economics 93, 1553–1590.

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

Bernanke, B.S., Gertler, M., 2001. *Should Central Banks Respond to Movements in Asset Prices?* American Economic Review 91, 253–257.

Bertsch, C., Hull, I., Lumsdaine, R.L., Zhang, X., 2025. *Central bank mandates and monetary policy stances: Through the lens of Federal Reserve speeches*. Journal of Econometrics 249, 105948.

Bischof, J.M., Airoldi, E.M., 2012. *Summarizing topical content with word frequency and exclusivity*, in: Proceedings of the 29th International Coference on Machine Learning, Omnipress, Madison, WI, USA. p. 9–16.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. *Latent dirichlet allocation*. Journal of machine Learning research 3, 993–1022.

Bohl, M.T., Kanelis, D., Siklos, P.L., 2023. *Central Bank Mandates: How Differences Can Influence the Content and Tone of Central Bank Communication*. Journal of International Money and Finance 130.

Born, B., Ehrmann, M., Fratzscher, M., 2014. *Central Bank Communication On Financial Stability*. The Economic Journal 124, 701–734.

Boyarchenko, N., Favara, G., Schularick, M., 2022. *Financial Stability Considerations for Monetary Policy: Empirical Evidence and Challenges*. Finance and Economics Discussion Series 2022-006. Washington: Board of Governors of the Federal Reserve System.

Carvalho, C., Nechio, F., Tristão, T., 2021. *Taylor rule estimation by OLS*. Journal of Monetary Economics 124, 140–154.

Chadha, J.S., Macchiarelli, C., Goel, S., Hantzsche, A., Mellina, S., 2026. *Deciphering Delphic Guidance: The Bank of England and Geopolitical Uncertainty*. Journal of International Money and Finance 162, 103530.

Cieslak, A., McMahon, M., 2024. *Tough Talk: The Fed and the Risk Premium*. SSRN .

Cohen, J., 1960. *A Coefficient of Agreement for Nominal Scales*. Educational and Psychological Measurement 20, 37–46.

Coibion, O., Gorodnichenko, Y., 2012. *Why Are Target Interest Rate Changes So Persistent?* American Economic Journal: Macroeconomics 4, 126–62.

Correa, R., Garud, K., Londono, J.M., Mislang, N., 2021. *Sentiment in Central Banks' Financial Stability Reports*. Review of Finance 25, 85–120.

Cukierman, A., 2009. *The Limits of Transparency*. Economic Notes 38, 1–37.

Danker, D.J., Luecke, M.M., 2005. *Background on FOMC Meeting Minutes*. Federal Reserve Bulletin 91, 175–179.

Devlin, J., Chang, M.W., Lee, K., Toutanova, K., 2019. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies .

Dybowski, T.P., Kempa, B., 2020. *The European Central Bank's monetary pillar after the financial crisis*. Journal of Banking and Finance 121, 105965.

Dyer, T., Lang, M., Stice-Lawrence, L., 2017. *The Evolution of 10-K Textual Disclosure: Evidence from Latent Dirichlet Allocation*. Journal of Accounting and Economics 64, 221–245.

Ehrmann, M., Fratzscher, M., 2007. *Communication by Central Bank Committee Members: Different Strategies, Same Effectiveness?* Journal of Money, Credit and Banking 39, 509–541.

Ferrara, F.M., Masciandaro, D., Moschella, M., Romelli, D., 2022. *Political voice on monetary policy: Evidence from the parliamentary hearings of the European Central Bank.* European Journal of Political Economy 74, 102143.

Grimm, M., Jordà, Ò., Schularick, M., Taylor, A.M., 2026. *Loose Monetary Policy and Financial Instability.* The Review of Economic Studies Forthcoming.

Hansen, S., McMahon, M., Prat, A., 2018. *Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach*.* The Quarterly Journal of Economics 133, 801–870.

Hernández-Murillo, R., Shell, H., 2014. *The rising complexity of the FOMC statement.* Economic Synopses 23.

Huang, A.H., Wang, H., Yang, Y., 2023. *FinBERT: A Large Language Model for Extracting Information from Financial Text*.* Contemporary Accounting Research 40, 806–841.

Istrefi, K., Odendahl, F., Sestieri, G., 2023. *Fed communication on financial stability concerns and monetary policy decisions: Revelations from speeches.* Journal of Banking and Finance 151, 106820.

Jegadeesh, N., Wu, D., 2017. *Deciphering Fedspeak: The Information Content of FOMC Meetings.* Working Paper .

Kanelis, D., Siklos, P.L., 2025. *The ECB press conference statement: deriving a new sentiment indicator for the euro area.* International Journal of Finance & Economics 30(1), 652–664.

Kashyap, A.K., Siegert, C., 2020. *Financial Stability Considerations and Monetary Policy.* International Journal of Central Banking 16, 231–266.

Kashyap, A.K., Stein, J.C., 2023. *Monetary Policy When the Central Bank Shapes Financial-Market Sentiment.* Journal of Economic Perspectives 37, 53–76.

Kazinnik, S., Sinclair, T.M., 2026. *FOMC In Silico: A Multi-Agent System for Monetary Policy Decision Modeling.* SSRN Working Paper .

Krippner, L., 2013. *Measuring the stance of monetary policy in zero lower bound environments.* Economics Letters 118, 135–138.

Landis, J.R., Koch, G.G., 1977. *An Application of Hierarchical Kappa-type Statistics in the Assessment of Majority Agreement among Multiple Observers.* Biometrics 33, 363–374.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V., 2019. *RoBERTa: A Robustly Optimized BERT Pretraining Approach* ArXiv: 1907.11692.

Loughran, T., McDonald, B., 2011. *When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks*. The Journal of Finance 66, 35–65.

Lundberg, S.M., Lee, S., 2017. *A unified approach to interpreting model predictions*. CoRR abs/1705.07874.

Maibaum, F., Kriebel, J., Foege, J.N., 2024. *Selecting textual analysis tools to classify sustainability information in corporate reporting*. Decision Support Systems 183, 114269.

Meade, E.E., Burk, N.A., Josselyn, M., 2015. *The FOMC Meeting Minutes: An Assessment of Counting Words and the Diversity of Views*. FEDS Notes, Board of Governors of the Federal Reserve System, May 26.

Mimno, D., Wallach, H., Talley, E., Leenders, M., McCallum, A., 2011. *Optimizing Semantic Coherence in Topic Models*. Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing , 262–272.

Mishev, K., Gjorgjevikj, A., Vodenska, I., Chitkushev, L.T., Trajanov, D., 2020. *Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers*. IEEE Access 8, 131662–131682.

Moschella, M., Pinto, L., 2019. *Central banks' communication as reputation management: How the Fed talks under uncertainty*. Public Administration 97, 513–529.

Oet, M.V., Lyytinen, K., 2017. *Does Financial Stability Matter to the Fed in Setting US Monetary Policy?* Review of Finance 21, 389–432.

Orphanides, A., 2001. *Monetary Policy Rules Based on Real-Time Data*. American Economic Review 91, 964–985.

Orphanides, A., 2024. *The Federal Reserve's Interpretation and Implementation of its Mandate*. Hoover Institution Working Paper .

Peek, J., Rosengren, E.S., Tootell, G.M.B., 2016. Does Fed policy reveal a ternary mandate? Federal Reserve Bank of Boston Working Papers .

Pfeifer, M., Marohl, V.P., 2023. *CentralBankRoBERTa: A fine-tuned large language model for central bank communications*. The Journal of Finance and Data Science 9, 100114.

Reinhart, C.M., Rogoff, K.S., 2009. *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press, Princeton, NJ.

Roberts, M.E., Stewart, B.M., Airoldi, E.M., 2016. *A Model of Text for Experimentation in the Social Sciences*. Journal of the American Statistical Association 111, 988–1003.

Roberts, M.E., Stewart, B.M., Tingley, D., 2019. *stm: An R Package for Structural Topic Models*. Journal of Statistical Software 91, 1–40.

Romer, C.D., Romer, D.H., 2023. *Presidential Address: Does Monetary Policy Matter? The Narrative Approach after 35 Years*. American Economic Review 113, 1395–1423.

Schinasi, G.J., 2003. *Responsibility of Central Banks for Stability in Financial Markets*. IMF Working Papers .

Shapiro, A.H., Wilson, D.J., 2022. *Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives using Text Analysis*. The Review of Economic Studies 89, 2768–2805.

Shapley, L.S., 1953. A value for n-person games, in: Kuhn, H.W., Tucker, A.W. (Eds.), Contributions to the Theory of Games, Volume II. Princeton University Press, Princeton, pp. 307–318.

Siklos, P., St. Amand, S., Wajda, J., 2018. *The Evolving Scope and Content of Central Bank Speeches*. CIGI Papers 202.

Smets, F., 2014. *Financial Stability and Monetary Policy: How Closely Interlinked?* International Journal of Central Banking 10, 263–300.

Stein, J.C., 2013. *Overheating in Credit Markets: Origins, Measurement, and Policy Responses*. Speech at "Restoring Household Financial Stability after the Great Recession: Why Household Balance Sheets Matter" A Research Symposium sponsored by the Federal Reserve Bank of St. Louis, St. Louis, Missouri.

Svensson, L.E., 2017. *Leaning Against the Wind: The Role of Different Assumptions About the Costs*. National Bureau of Economic Research Working Paper 23745.

Svensson, L.E.O., 2003. *Monetary Policy and Real Stabilization*. National Bureau of Economic Research Working Paper Series Working Paper No. 9486.

Tadle, R.C., 2022. *FOMC minutes sentiments and their impact on financial markets*. Journal of Economics and Business 118, 106021.

Tarullo, D.K., 2014. Monetary policy and financial stability. *Speech at the 30th Annual National Association for Business Economics Economic Policy Conference, Arlington, Virginia*.

Tetlock, P.C., 2007. *Giving Content to Investor Sentiment: The Role of Media in the Stock Market*. The Journal of Finance 62, 1139–1168.

Wischnewsky, A., Jansen, D.J., Neuenkirch, M., 2021. *Financial stability and the Fed: Evidence from congressional hearings*. Economic Inquiry 59, 1192–1214.

Wu, J.C., Xia, F.D., 2016. *Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound*. Journal of Money, Credit and Banking 48, 253–291.

Yellen, J.L., 2013. *Challenges Confronting Monetary Policy.* Speech at the National Association for Business Economics Policy Conference.

Yellen, J.L., 2014. *Monetary Policy and Financial Stability.* Speech at the 2014 Michel Camdessus Central Banking Lecture, International Monetary Fund, Washington, D.C.

Appendix to "The Financial Instability - Monetary Policy Nexus: Evidence from the FOMC Minutes"

Dimitrios Kanelis Lars H. Kranzmann Pierre L. Siklos

A. The Methodology Behind Structural Topic Modeling

A.1. Structural Topic Modeling and Topic Identification

To identify the latent topics discussed in FOMC deliberations, we employ the STM developed by Roberts et al. (2016). STM is an unsupervised machine learning approach that identifies topics as distributions over words and documents as mixtures of topics. In our setting, we feed the paragraphs of the FOMC minutes as input data into the model. This choice is motivated by the highly structured nature of the minutes, where individual paragraphs typically focus on a single issue or policy consideration. Using paragraphs rather than entire meeting minutes therefore allows the STM to recover more clearly defined and economically meaningful topics.

Prior to estimating the model, we preprocess the text by converting all words to lowercase, removing punctuation and stop words, and reducing words to their stems. We then estimate a series of STM specifications that differ in the number of topics, K.

Choosing the appropriate number of topics is a central challenge in topic modeling. A larger number of topics generally leads to more exclusive and differentiated topics, while a smaller number of topics tends to increase semantic coherence by grouping related concepts together. As discussed by Roberts et al. (2019), there is no universally optimal choice of K, as topic exclusivity and semantic coherence involve an inherent trade-off. Following the literature on topic modeling, we evaluate alternative specifications based on the criteria of *topic exclusivity* and *semantic coherence*. Specifically, we estimate models with $K \in [4, 60]$ and compute topic exclusivity (Bischof and Airoldi, 2012) and semantic coherence (Mimno et al., 2011) for each specification.

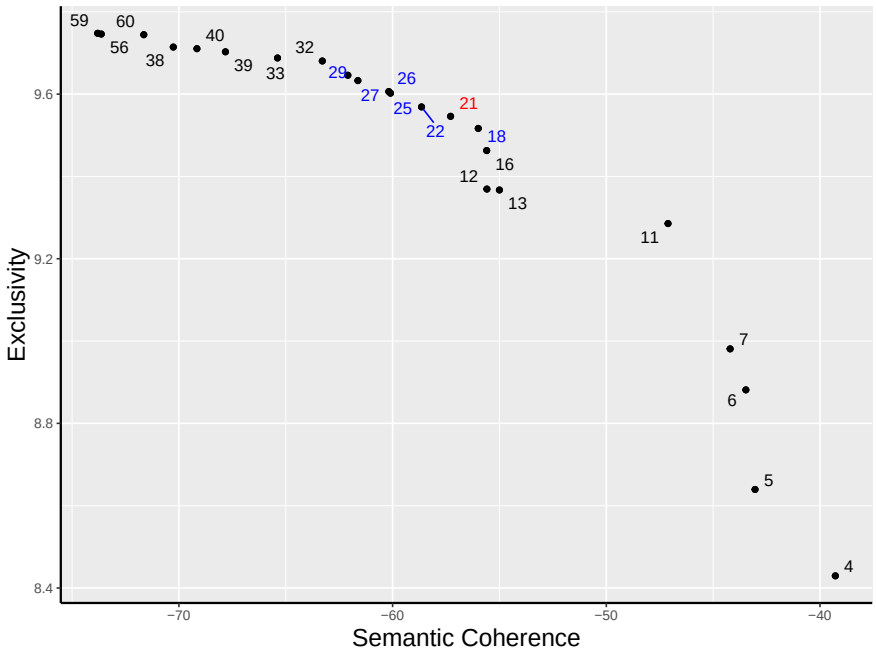
Figure A1 reports the resulting exclusivity–coherence frontier. Consistent with previous applications of STM (Bohl et al., 2023), models with lower K generate more semantically coherent but substantially less differentiated topics, whereas models with higher K improve topic exclusivity at the expense of coherence. We therefore focus on specifications that lie on the Pareto frontier and are not strictly dominated in either dimension.

For the baseline analysis, we select $K = 21$. This specification provides the most meaningful balance between semantic coherence and topic exclusivity and yields economically interpretable topics that can be clearly distinguished from one another. However, to ensure that our findings do not depend on a particular choice of K, we conduct in Appendix B.2 extensive robustness checks and sensitivity analysis using different values of K. Following the evidence from the exclusivity–coherence analysis, we focus on models with topic numbers between 16 and 29, thereby avoiding specifications that have either insufficient topic differentiation or very low semantic coherence.

After selecting the baseline specification, we inspect both the highest-probability word stems and representative high-gamma paragraphs for each topic. We provide the representative topics for *Financial Stability* in Table A1 and for *Mandate Discussions* in Table A2. Based on this combined quantitative and qualitative assessment, we assign economically meaningful

labels to the resulting topics. Table 1 reports the complete list of topic labels together with their most influential word stems.

Figure A1: Pareto-efficient Topic Frontier



Note: This figure illustrates the trade-off between exclusivity and semantic coherence for the discussion corpus of the minutes. We estimate STM for $K \in \{4, \ldots, 60\}$ and remove all models strictly dominated in metrics. We focus on the topic numbers $K \in \{18, 21, 22, 25, 26, 27, 29\}$ as possible classifiers based on interpretability.

Table A1: Financial Stability: Representative Topics

Topic	Gamma	Date	Paragraph
2	0.64	2009-04-29	In these circumstances, the Federal Reserve will employ all available tools to promote economic recovery and to preserve price stability. The Committee will maintain the target range for the federal funds rate at 0 to 1/4 percent and anticipates that economic conditions are likely to warrant exceptionally low levels of the federal funds rate for an extended period. As previously announced, to provide support to mortgage lending and housing markets and to improve overall conditions in private credit markets, the Federal Reserve will purchase a total of up to $1.25 trillion of agency mortgage-backed securities and up to $200 billion of agency debt by the end of the year. In addition, the Federal Reserve will buy up to $300 billion of Treasury securities by autumn. The Committee will continue to evaluate the timing and overall amounts of its purchases of securities in light of the evolving economic outlook and conditions in financial markets. The Federal Reserve is facilitating the extension of credit to households and businesses and supporting the functioning of financial markets through a range of liquidity programs. The Committee will continue to carefully monitor the size and composition of the Federal Reserve's balance sheet in light of financial and economic developments.
2	0.92	2011-09-21	Most members also supported a change in the Committee's reinvestment policy. To help support conditions in mortgage markets, the Committee decided to reinvest principal received from its holdings of agency debt and agency MBS in agency MBS rather than continuing to reinvest in longer-term Treasury securities as had been the Committee's practice for more than a year. The effect of this change will be to keep the SOMA's holdings of agency securities at an approximately constant level; under the previous practice, those holdings were declining on an ongoing basis. This change in reinvestment policy was expected to help reduce the spread between yields on mortgage-backed securities and those on comparable-maturity Treasury securities seen this year and so contribute to lower mortgage rates. Members also noted that the change in reinvestment policy could help prevent the shares of outstanding longer-term Treasury securities held by the Federal Reserve from reaching levels high enough to result in a deterioration in Treasury market functioning. One member who opposed the maturity extension program also opposed the change in reinvestment policy because he judged that it would not benefit housing markets. At the same time, the Committee decided to maintain its existing policy of rolling over maturing Treasury securities at auction.
2	0.93	2015-10-28	The Committee also maintained its policy of reinvesting principal payments from its agency debt and agency mortgage-backed securities in agency mortgage-backed securities and of rolling over maturing Treasury securities at auction. This policy, by keeping the Committee's holdings of longer-term securities at sizable levels, should help maintain accommodative financial conditions.
2	0.94	2016-03-16	The Committee also decided to maintain its existing policy of reinvesting principal payments from its holdings of agency debt and agency mortgage-backed securities in agency mortgage-backed securities and of rolling over maturing Treasury securities at auction, and it anticipated doing so until normalization of the level of the federal funds rate is well under way. This policy, by keeping the Committee's holdings of longer-term securities at sizable levels, should help maintain accommodative financial conditions.
2	0.72	2020-03-15	Participants also considered open market operations to purchase Treasury securities and agency MBS to support the smooth functioning of these securities markets, which in turn would help support the supply of credit to households and businesses. Participants generally agreed that, over the coming months, it would be appropriate to increase the Federal Reserve's holdings of Treasury securities by at least $500 billion and its holdings of agency MBS by at least $200 billion. Additionally, all principal payments from the Federal Reserve's holdings of agency debt and agency MBS would be reinvested in agency MBS. Those Treasury and agency MBS purchases would be in addition to the recently expanded overnight and term repo operations conducted by the Desk. Participants stressed that it was important to communicate that the Committee would be prepared to increase the size of the securities purchases, as needed, on the basis of its close monitoring of market conditions. Some participants noted that it was important to stress in communications that the primary purpose of these asset purchases was to support the smooth functioning of Treasury and agency MBS markets rather than to provide further monetary policy accommodation by pushing down longer-term yields. A couple of participants noted that because some of the purchases would be at longer maturities, the purchases could provide some accommodation by lowering longer-term yields.
2	0.93	2020-04-29	To support the flow of credit to households and businesses, members agreed that it was appropriate for the Federal Reserve to continue to purchase Treasury securities and agency RMBS and CMBS in the amounts needed to support smooth market functioning, thereby fostering effective transmission of monetary policy to broader financial conditions. In addition, the Desk would continue to offer large-scale overnight and term repo operations. Members agreed that they would closely monitor market conditions and be prepared to adjust their plans as appropriate.

(continued)

Topic	Gamma	Date	Paragraph
3	0.82	2007-12-11	Participants discussed in detail the resurgence of stresses in financial markets in November. The renewed stresses reflected evidence that the performance of mortgage-related assets was deteriorating further, potentially increasing the losses that were being borne in part by a number of major financial firms, including money-center banks, housing-related government-sponsored enterprises, investment banks, and financial guarantors. Moreover, participants recognized that some lenders might be exposed to additional losses: Delinquency rates on credit card loans, auto loans, and other forms of consumer credit, while still moderate, had increased somewhat, particularly in areas hard hit by house price declines and mortgage defaults. Past and prospective losses appeared to be spurring lenders to tighten further the terms on new extensions of credit, not just in the troubled markets for nonconforming mortgages but, in some cases, for other forms of credit as well. In addition, participants noted that some intermediaries were facing balance sheet pressures and could become constrained by concerns about rating-agency or regulatory capital requirements. Among other factors, banks were experiencing unanticipated growth in loans as a result of continuing illiquidity in the market for leveraged loans, persisting problems in the commercial paper market that had sparked draws on back-up lines of credit, and more recently, consolidation of assets of off-balance-sheet affiliates onto banks' balance sheets.
3	0.90	2008-10-29	Participants noted that financial conditions had worsened significantly over the intermeeting period. The failure or near failure of a number of major financial institutions had deepened market concerns about counterparty credit risk and liquidity risk. As a result, financial intermediaries had cut back on lending to some counterparties, particularly for terms beyond overnight, and in general were conserving liquidity and capital. Moreover, risk aversion of investors increased, driving credit spreads sharply higher. Survey results and anecdotal information also suggested that credit conditions had tightened significantly further for businesses and households. Equity prices had varied widely and were substantially lower, on net. Participants saw the potential for financial strains to intensify if some investors, such as hedge funds, found it necessary to sell assets and as lending institutions built reserves against losses. Participants were concerned that the negative spiral in which financial strains lead to weaker spending, which in turn leads to higher loan losses and a further deterioration in financial conditions, could persist for a while longer. While the global efforts to recapitalize banks and guarantee deposits had helped stabilize the situation, risk spreads remained higher, asset prices lower, and credit conditions tighter than prior to the recent disruptions. Moreover, some participants noted that the specifics and effectiveness of some government programs to support financial markets and institutions remained unclear.
3	0.81	2011-06-22	Changes in financial market conditions since the April meeting suggested that investors had become more concerned about risk. Equity markets had seen a broad selloff, and risk spreads for many corporate borrowers had widened noticeably. Large businesses that have access to capital markets continued to enjoy ready access to credit–including syndicated loans–on relatively attractive terms; however, credit conditions remained tight for smaller, bank-dependent firms. Bankers again reported gradual improvements in credit quality and generally weak loan demand. In identifying possible risks to financial stability, a few participants expressed concern that credit conditions in some sectors–most notably the agriculture sector–might have eased too much amid signs that investors in these markets were aggressively taking on more leverage and risk in order to obtain higher returns. Meeting participants also noted that an escalation of the fiscal difficulties in Greece and spreading concerns about other peripheral European countries could cause significant financial strains in the United States. It was pointed out that some U.S. money market mutual funds have significant exposures to financial institutions from core European countries, which, in turn, have substantial exposures to Greek sovereign debt. Participants were also concerned about the possible effect on financial markets of a failure to raise the statutory federal debt ceiling in a timely manner. While admitting that it was difficult to know what the precise effects of such a development would be, participants emphasized that even a short delay in the payment of principal or interest on the Treasury Department's debt obligations would likely cause severe market disruptions and could also have a lasting effect on U.S. borrowing costs.
3	0.90	2014-04-30	In their discussion of financial stability, participants generally did not see imbalances that posed significant near-term risks to the financial system and the broader economy, but they nevertheless reviewed some financial developments that pointed to potential future risks. A couple of participants noted that conditions in the leveraged loan market had become stretched, although equity cushions on new deals remained above levels seen prior to the financial crisis. Two others saw declining credit spreads, particularly on speculative-grade corporate bonds, as consistent with an increase in investors' appetite for risk. In addition, several participants noted that the low level of expected volatility implied by some financial market prices might also signal an increase in risk appetite. Some stated that it would be helpful to continue to explore the appropriate regulatory, supervisory, and monetary policy responses to potential risks to financial stability.

(continued)

Topic	Gamma	Date	Paragraph
3	0.90	2017-05-03	With regard to financial stability, several participants emphasized that higher requirements for capital and liquidity in the banking system and other prudential standards had contributed to increased resilience in the financial system since the financial crisis. However, they expressed concerns that a possible easing of regulatory standards could increase risks to financial stability. In addition, it was noted that real estate values were elevated in some sectors of the CRE market, that a sharp decline in such valuations could pose risks to financial stability, and that potential reforms in the housing finance sector could have implications for such valuations.
9	0.72	2001-06-27	Housing activity remained at a high level as attractive mortgage interest rates evidently continued to counterbalance the negative effects on consumer attitudes of somewhat weaker labor markets and reduced stock market wealth. While housing activity in a number of areas continued to be described as fairly robust, members noted that residential sales and construction had slipped in some parts of the nation. Even so, given existing backlogs and the continued availability of attractive mortgage rates, nationwide housing construction was expected to remain near its currently elevated level.
9	0.74	2007-06-28	Participants generally agreed that the housing sector was likely to remain a drag on growth for some time yet and represented the most significant downside risk to the economic outlook. Although starts of single-family homes had moved up, on balance, over recent months, permits for new construction continued to decline. A number of participants noted that inventories of new homes for sale remained quite elevated. Housing activity was seen as likely to continue to contract for several more quarters. Participants also identified a number of downside risks associated with their outlook for residential construction. The recent increase in interest rates for prime mortgages could further dampen the demand for housing. Moreover, a number of participants pointed to rising mortgage delinquency rates and related difficulties in the subprime mortgage market as factors that could crimp the availability of mortgage credit and the demand for housing.
9	0.93	2012-04-25	Recent housing-sector indicators, including sales and starts, suggested some upward movement, but some participants saw the improvement as likely related to unusually warm winter weather in much of the country. Overall, the level of activity in the sector remained depressed. House prices appeared to be stabilizing but had not yet begun to rise in most markets. Most participants anticipated that the housing sector was likely to recover only slowly over time, but a few were more optimistic about the potential for a more rapid housing recovery given reports of stronger demand in some regions and of improved sentiment among builders, as well as signs that recent changes to the Home Affordable Refinance Program were contributing to the refinancing of performing high loan-to-value mortgages.
9	0.80	2014-09-17	The recovery in housing activity remained slow according to most participants. Although mortgage rates were still low and housing appeared to be relatively affordable, various factors were seen as restraining demand, including low expected income and high levels of student debt as well as difficulty in obtaining mortgage credit, particularly for younger, first-time homebuyers. It was also noted that the weakness in homebuilding along with the continued rise in house prices suggested that supply constraints were also weighing on construction activity. A couple of participants indicated that some demand appeared to have shifted to rental properties. The rising demand for rentals was in part being satisfied by investors buying homes for the rental market; it was also providing support for multifamily construction. Some participants noted their concern that a number of the factors restraining residential construction might persist, damping the housing recovery for some time.
9	0.80	2015-09-17	Housing activity was improving, with sales and new construction trending higher. Solid gains in employment and favorable mortgage rates were anticipated to continue to underpin the recovery in housing. Contacts in a number of Districts were upbeat about prospects for the sector, citing strengthening sales, rising home prices, an upturn in household formations, and reports that buyers had accelerated purchases in anticipation of the possibility that mortgage rates might move higher in the near term. Multifamily construction was particularly strong in a couple of Districts, but in another a shortage of lots was constraining builders' ability to meet strong demand for new single-family homes.

Note: This table presents paragraphs categorized under financial stability topics, identified using an STM with 21 topics. An overview of all 21 topics is presented in Table 1. The gamma value represents the probability that a paragraph belongs to a specific topic. Paragraphs are sorted by the topic with the highest gamma value.

Table A2: Mandate: Representative Topics

Topic	Gamma	Date	Paragraph
4	0.61	2006-09-20	Business investment spending generally was seen as expanding at a reasonably good pace. Meeting participants noted broad strength in manufacturing of capital goods. Nonresidential construction activity continued to strengthen, and in the process was absorbing some of the resources that were no longer employed in homebuilding. Although some survey evidence suggested that some firms were trimming capital spending plans, participants reported that their business contacts generally were quite positive about the economic outlook and the strength of demand for their products. In this environment, investment spending would likely continue to be supported by expansion of overall output, strong balance sheets and profits, and the ready availability of funding from financial markets and institutions.
4	0.56	2013-07-31	In the business sector, the outlook still appeared to be mixed. Manufacturing activity was reported to have picked up in a number of Districts, and activity in the energy sector remained at a high level. Although a step-up in business investment was likely to be a necessary element of the projected pickup in economic growth, reports from businesses ranged from those contacts who expressed heightened optimism to those who suggested that little acceleration was likely in the second half of the year.
4	0.81	2003-05-06	Business fixed investment remained a key factor in the prospects for overall economic activity, and persisting weakness in such spending in association with gloomy sentiment and a high degree of risk aversion among business decisionmakers did not bode well for the capital investment outlook, at least for the near term. Anecdotal reports by business contacts tended to emphasize widespread excess capacity as a reason for holding down business capital spending, including high vacancy rates in office and other business structures. In this atmosphere, most business decisionmakers evidently preferred to rely on the increasingly efficient or fuller utilization of existing producer facilities rather than expanding the latter to meet growth in demand. Indeed, according to business contacts, investment expenditures generally were limited to replacement and to some extent to upgrading of existing facilities rather than for expansion. In some cases, businesses reportedly were acquiring used capital equipment and unoccupied building space at greatly reduced costs, thereby holding down the current production of new capital but also relieving selling firms of some excess capacity.
5	0.85	2016-01-27	A broad range of indicators showed ongoing improvement in labor market conditions. Most notably, increases in nonfarm payroll employment were quite strong during the final three months of 2015. Although the unemployment rate, at 5.0 percent, was unchanged over that period, it was at a level close to or below most participants' estimates of its longer-run normal rate. Moreover, the labor force participation rate and the employment-to-population rate moved up toward year-end. Many viewed labor market underutilization as having been substantially reduced over the past year, and a few saw slack as having been largely eliminated. In their comments on labor market conditions, participants cited strong employment gains, low levels of unemployment in their Districts, reports of shortages of workers in various industries, or firming in wage increases. Most anticipated that employment would expand at a solid rate over the year ahead, although several saw the prospect of some moderation in employment gains from the particularly large increases in the fourth quarter of 2015.
5	0.77	2004-05-04	After a protracted period of meager gains in employment, conditions in the U.S. labor market evidently were improving in recent weeks. In addition to noting the substantial jump in payrolls in March, several members relayed anecdotal information from business contacts around the nation that hiring was continuing to pick up and that firms were planning further increases in workforces. Some temporary help firms reported rising demand, a possible precursor of a pickup in permanent hiring. A number of members cited reports of difficulties in hiring within certain job families in which specialized skills were in short supply together with indications that wage increases in those occupations tended to be larger than average. Even so, considerable slack seemed to remain in the labor market overall, and wage gains on the whole were moderate.

(continued)

Topic	Gamma	Date	Paragraph
5	0.97	2011-12-13	A number of recent indicators showed some improvement in labor market conditions: Payroll employment had posted moderate gains for five months, new claims for unemployment insurance had drifted lower, and the unemployment rate had turned down. One participant noted that the series of upward revisions to the initial estimates of payroll employment in recent months was an encouraging sign of sustained hiring, although several participants remarked that they saw the labor market as still improving only slowly. Others indicated that because part of the recent decline in the jobless rate was associated with a reduction in labor force participation, the drop in the unemployment rate likely overstated the overall improvement in the labor market. Moreover, unemployment, particularly longer-term unemployment, remained high, and the number of involuntary part-time workers was still elevated. Some participants again expressed concern that the persistence of high levels of long-duration unemployment and the underutilization of the workforce could eventually lead to a loss of skills and an erosion of potential output. Another participant suggested that the unemployment rate was a more useful indicator of cyclical labor market developments than the level of employment relative to the size of the population, which was more likely to be influenced by structural changes in labor demand and supply. Participants expressed a range of views on the current extent of slack in the labor market. It was noted that because of factors including ongoing changes in the composition of available jobs and workers' skills, some part of the increase in unemployment since the beginning of the recession had been structural rather than cyclical. Others pointed out that the very modest increases in labor compensation of late suggested that underutilization of labor was still significant.
6	0.70	2017-11-01	In view of the persistent shortfall of inflation from the Committee's 2 percent objective and questions about whether longer-term inflation expectations were consistent with achievement of that objective, a couple of participants discussed the possibility that potential alternative frameworks for the conduct of monetary policy could be helpful in fulfilling the Committee's statutory mandate. One question, for example, was whether a framework that generally sought to keep the price level close to a gradually rising path–rather than the current approach in which the Committee does not seek to make up for past deviations of inflation from the 2 percent goal–might be more effective in fostering the Committee's objectives if the neutral level of the federal funds rate remains low.
6	0.53	2021-12-15	In support of the Committee's goals of maximum employment and inflation at the rate of 2 percent over the longer run, participants judged that it would be appropriate for the Committee to keep the target range for the federal funds rate at 0 to 1/4 percent until labor market conditions had reached levels consistent with the Committee's assessments of maximum employment, a condition most participants judged could be met relatively soon if the recent pace of labor market improvements continued. A few participants remarked that maximum employment consistent with price stability evolves over time and that further improvements in labor markets were likely over subsequent years as the economy continued to expand. Some participants also remarked that there could be circumstances in which it would be appropriate for the Committee to raise the target range for the federal funds rate before maximum employment had been fully achieved—for example, if the Committee judged that its employment and price-stability goals were not complementary in light of economic developments and that inflation pressures and inflation expectations were moving materially and persistently higher in a way that could impede the attainment of the Committee's longer-run goals.
6	0.95	2021-07-28	In their consideration of the stance of monetary policy, participants reaffirmed the Federal Reserve's commitment to using its full range of tools to support the U.S. economy during this challenging time, thereby promoting the Committee's statutory goals of maximum employment and price stability. Participants judged that the current stance of monetary policy remained appropriate to promote maximum employment as well as to achieve inflation that averages 2 percent over time and longer-term inflation expectations that are well anchored at 2 percent. Participants also reiterated that the existing outcome-based guidance implied that the paths of the federal funds rate and the balance sheet would depend on actual progress toward reaching the Committee's maximum-employment and inflation goals.

(continued)

Topic	Gamma	Date	Paragraph
10	0.58	2011-03-15	Participants expected that the boost to headline inflation from recent increases in energy and other commodity prices would be transitory and that underlying inflation trends would be little affected as long as commodity prices did not continue to rise rapidly and longer-term inflation expectations remained stable. However, a significant increase in longer-term inflation expectations could contribute to excessive wage and price inflation, which would be costly to eradicate. Accordingly, participants considered it important to pay close attention to the evolution not only of headline and core inflation but also of inflation expectations. In this regard, participants observed that measures of longer-term inflation compensation derived from financial instruments had remained stable of late, suggesting that longer-term inflation expectations had not changed appreciably, although measures of one-year inflation compensation had risen notably. Survey-based measures of inflation expectations also indicated that longer-term expected inflation had risen much less than near-term inflation expectations. A few participants noted that the adoption by the Committee of an explicit numerical inflation objective could help keep longer-term inflation expectations well anchored.
10	0.70	2015-04-29	In their discussion of the economic situation and the outlook, meeting participants regarded the information received over the intermeeting period as suggesting that economic growth had slowed during the winter months, in part reflecting transitory factors. The pace of job gains had moderated, and the unemployment rate had remained steady, with a range of labor market indicators suggesting that underutilization of labor resources was little changed. Most participants expected that, following the slowdown in the first quarter, real economic activity would resume expansion at a moderate pace, and that labor market conditions would improve further. Inflation continued to run below the Committee's longer-run objective, partly reflecting earlier declines in energy prices and decreasing prices of non-energy imports. Market-based measures of inflation compensation remained low, while survey-based measures of longer-term inflation expectations had remained stable. Participants generally anticipated that inflation would rise gradually toward the Committee's 2 percent objective as the labor market improved further and the transitory effects of declines in energy prices and non-energy import prices dissipated. Participants judged that recent domestic economic developments had increased uncertainty regarding the economic outlook. While participants continued to see potential downside risks resulting from foreign economic and financial developments, most still viewed the risks to the outlook for economic growth and the labor market as nearly balanced.
10	0.95	2017-05-03	Inflation, measured as the 12-month change in the headline PCE price index, had been running close to the Committee's 2 percent longer-run objective. Core inflation continued to run somewhat below 2 percent. Both headline and core consumer price indexes fell in March. Market-based measures of inflation compensation had remained low, while survey-based measures of longer-term inflation expectations had changed little on balance.
12	0.72	2003-10-28	The recent strength in final sales was associated with sizable inventory liquidation by business firms, and recent surveys and anecdotal commentary suggested that inventories were at unusually low levels in relation to sales, notably in manufacturing. In the circumstances, a continuation of robust final demand could be expected to foster efforts to rebuild inventories, with potentially substantial short-run stimulus to the economy. However, the timing and extent of such restocking were subject to uncertainty, and for now available reports indicated that business firms were continuing to follow a highly cautious approach to inventory investment.
12	0.52	2003-06-25	In the Committee's discussion of current and prospective economic developments, members referred to signs of improvement in some sectors of the economy, but they saw no conclusive evidence of an appreciable overall strengthening in the sluggish economic expansion. On the positive side, they pointed to reports of some pickup in retail sales, indications that labor and product markets might be stabilizing, continued robust activity in housing markets, and ongoing impetus from the federal government sector. Concurrently, however, weakness persisted in business capital expenditures, which members continued to view as the critical factor inhibiting the economic expansion. Looking ahead, they emphasized that favorable underlying conditions were in place to support a substantial acceleration of the expansion, though the timing and dimensions of a significantly improved economic performance remained uncertain. Positive factors bearing on the outlook mentioned by members included the accommodative stance of monetary policy and supportive financial conditions more generally, the persistence of rapid growth in labor productivity, sizable declines in energy prices from elevated levels earlier in the year, and indications of rising consumer confidence and of less negative business sentiment. Members also gave considerable emphasis to the anticipated effects of recent legislation that in short order would add substantially to the degree of fiscal stimulus.

(continued)

Topic	Gamma	Date	Paragraph
12	0.87	1995-11-15	Ongoing efforts by many business firms to bring inventories into better alignment with sales had resulted in declining inventory investment since earlier in the year. Some further inventory adjustments, notably in stocks of motor vehicles, were expected over coming months, though not at a pace that would have a marked retarding effect on economic activity. Over much of 1996, inventory investment was projected to be a more neutral factor in the economy, with accumulation proceeding at a pace in line with growth in final sales, but the risks of unexpected developments in this sector of the economy were always substantial.
16	0.59	2018-09-26	Meeting participants noted that a number of communities suffered devastating losses associated with Hurricane Florence. Despite the magnitude of the storm-related destruction, participants expected the imprint on the level of overall economic activity at the national level to be relatively modest, consistent with the experience following several previous major storms.
16	0.53	2018-08-01	Participants commented on a number of risks and uncertainties associated with their outlook for economic activity, the labor market, and inflation over the medium term. They generally continued to see fiscal policy and the strengthening of the labor market as supportive of economic growth in the near term. Some noted larger or more persistent positive effects of these factors as an upside risk to the outlook. A few participants indicated, however, that a faster-than-expected fading of the fiscal impetus or a greater-than-anticipated subsequent fiscal tightening constituted a downside risk. In addition, all participants pointed to ongoing trade disagreements and proposed trade measures as an important source of uncertainty and risks. Participants observed that if a large-scale and prolonged dispute over trade policies developed, there would likely be adverse effects on business sentiment, investment spending, and employment. Moreover, wide-ranging tariff increases would also reduce the purchasing power of U.S. households. Further negative effects in such a scenario could include reductions in productivity and disruptions of supply chains. Other downside risks cited included the possibility of a significant weakening in the housing sector, a sharp increase in oil prices, or a severe slowdown in EMEs.
16	0.95	2021-06-16	In their discussion of current conditions, participants noted that progress on vaccinations had reduced the spread of COVID-19 in the United States. Amid this progress and strong policy support, indicators of economic activity and employment had strengthened. The sectors most adversely affected by the pandemic remained weak but had shown improvement. Inflation had risen, largely reflecting transitory factors. Overall financial conditions remained accommodative, in part reflecting policy measures to support the economy and the flow of credit to U.S. households and businesses. Participants generally noted that the path of the economy would depend significantly on the course of the virus. Progress on vaccinations would likely continue to reduce the effects of the public health crisis on the economy, but risks to the economic outlook remained.
18	0.88	2013-03-15	After assessing current conditions and the outlook for economic activity, the labor market, and inflation, all but one member agreed to raise the target range for the federal funds rate to 3/4 to 1 percent. This increase was viewed as appropriate in light of the further progress that had been made toward the Committee's objectives of maximum employment and 2 percent inflation. Members generally noted that the increase in the target range did not reflect changes in their assessments of the economic outlook or the appropriate path of the federal funds rate, adding that the increase was consistent with the gradual pace of removal of accommodation that was anticipated in December, when the Committee last raised the target range.
18	0.64	2015-09-17	In assessing whether economic conditions had improved sufficiently to initiate a firming in the stance of policy, many members said that the improvement in labor market conditions met or would soon meet one of the Committee's criteria for beginning policy normalization. But some indicated that their confidence that inflation would gradually return to the Committee's 2 percent objective over the medium term had not increased, in large part because recent global economic and financial developments had imparted some restraint to the economic outlook and placed further downward pressure on inflation in the near term. Most members agreed that their confidence that inflation would move to the Committee's inflation objective would increase if, as expected, economic activity continued to expand at a moderate rate and labor market conditions improved further. Many expected those conditions to be met later this year, although several members were concerned about downside risks to the outlook for real activity and inflation.

(continued)

Topic	Gamma	Date	Paragraph
18	0.97	2018-11-08	After assessing current conditions and the outlook for economic activity, the labor market, and inflation, members decided to maintain the target range for the federal funds rate at 2 to 2-1/4 percent. Members agreed that the timing and size of future adjustments to the target range for the federal funds rate would depend on their assessment of realized and expected economic conditions relative to the Committee's maximum employment and symmetric 2 percent inflation objectives. They reiterated that this assessment would take into account a wide range of information, including measures of labor market conditions, indicators of inflation pressures and inflation expectations, and readings on financial and international developments. More generally, members noted that decisions regarding near-term adjustments of the stance of monetary policy would appropriately remain dependent on the evolution of the outlook as informed by incoming data.
19	0.68	2004-01-28	In the Committee's discussion of policy for the intermeeting period ahead, all the members favored an unchanged policy stance that was directed toward maintaining reserve conditions consistent with a target federal funds rate of 1 percent. While the members were persuaded that a relatively vigorous economic expansion was now firmly established and was likely at some point to call for a move toward a more neutral policy stance, they concluded that such an adjustment was not warranted under current circumstances. In this regard they stressed that unused labor and other resources remained substantial, that inflation was at a very low level, and that inflation was not expected to change appreciably in either direction over the year ahead. Members acknowledged that there were risks in maintaining what might eventually prove to be an overly accommodative policy stance, but for now they judged that it was desirable to take risks on the side of assuring the rapid elimination of economic slack.
19	0.55	2003-01-29	In the Committee's discussion of policy for the intermeeting period, all the members supported a proposal to maintain an unchanged policy stance. While the economy had continued to grow slowly, monetary policy and overall financial conditions had remained accommodative and the prospects for an appreciable strengthening of the economic expansion over time were favorable. As some of the prevailing uncertainties currently impairing spending began to lift, possibly in the near term with regard to military developments in the Middle East, the Committee should be in a much better position to assess the underlying strength of the economy and the appropriate policy response. At this point, the Committee could not rule out a range of plausible economic outcomes, including the possibility of a persisting subpar economic performance or a much stronger than forecast acceleration of the expansion. Indeed, the Committee could envision circumstances when it might find it desirable to adjust its policy stance substantially and promptly in one direction or the other in the months ahead. The members concluded that a wait-and-see policy stance was desirable pending an improved basis for judging the ongoing performance of the economy. They also agreed that the accommodative stance of policy, developments over the intermeeting period, and their current forecasts in the context of tensions abroad argued for retaining a balanced risks assessment to be included in the statement that would be made public shortly after this meeting.
19	0.86	1999-02-03	In light of the uncertainties and diversity of risks surrounding the economic outlook, most members were in favor of retaining the existing symmetry of the directive. In one view, however, the risks of rising inflation were strong enough to warrant consideration of an asymmetrical directive that was tilted toward restraint. Nonetheless, since inflation was difficult to predict and any needed adjustment to policy in the period ahead could readily be implemented even with a symmetrical directive, all the members indicated that they could accept such a directive.
20	0.64	2011-03-15	In contrast to headline inflation, core inflation and other measures of underlying inflation remained subdued, though they appeared to have bottomed out. A number of participants noted that, with significant slack in resource utilization and with longer-term inflation expectations stable, underlying inflation likely would remain subdued for some time. However, the importance of resource slack as a factor influencing inflation was debated. Some participants pointed to research indicating that measures of slack were useful in predicting inflation. Others argued that, historically, such measures were only modestly helpful in explaining large movements in inflation; one noted the 2003-04 episode in which core inflation rose rapidly over a few quarters even though there appeared to be substantial resource slack.

(continued)

Topic	Gamma	Date	Paragraph
20	0.53	1993-11-16	With regard to the outlook for inflation over the year ahead, views did not vary greatly among the members. They ranged from expectations of some limited progress toward price stability to forecasts of a marginal increase in the core rate of inflation. Members who anticipated a relatively favorable inflation performance tended to underscore the likely persistence of appreciable slack in labor and other production resources on the assumption that growth in overall economic activity would remain on a moderate trend in line with their forecasts. Some also pointed to the absence of inflationary pressures in most commodity markets, the persistence of intense competition in local markets across the nation, and the outlook for relatively subdued increases in labor costs in part because of ongoing improvements in productivity. Other members gave more emphasis to the possibility that the economic expansion next year, especially if it turned out on the high side of the range encompassing the members' current projections, was more likely to be associated with some upward pressures on costs and prices. In this connection, relatively rapid growth in economic activity, should it persist into the early part of next year, probably would trigger attempts to raise prices and wages somewhat more rapidly even in the context of some continuing slack in overall capacity and labor utilization. At this point, however, there were no significant indications of accelerating inflation and business contacts around the nation did not currently see or seem to anticipate increasing inflationary pressures.
20	0.86	2000-02-02	Concerning the outlook for inflation, the members continued to see the risks as primarily tilted toward rising inflationary pressures, though they anticipated that further gains in productivity would hold down increases in unit labor costs and prices, at least over the nearer term. A key issue was whether growth in aggregate demand would moderate sufficiently to at least avoid greater pressures on what were already very tight labor markets. In this regard, several cited recent statistical and anecdotal evidence of larger increases in labor compensation, although unit labor costs did not appear to be trending higher at this point. However, some nonlabor input prices already were rising faster. The prospects for energy prices were very difficult to predict, but even if such prices were to stabilize, the passthrough of the large earlier increases into inflation and wage expectations, as well as into the prices of products that were heavily energy dependent, was likely to exert some upward pressure on prices throughout the economy.
21	0.65	2009-12-16	In their discussion of the economic situation and outlook, meeting participants agreed that the incoming data and information received from business contacts suggested that economic growth was strengthening in the fourth quarter, that firms were reducing payrolls at a less rapid pace, and that downside risks to the outlook for economic growth had diminished a bit further. Although some of the recent data had been better than anticipated, most participants saw the incoming information as broadly in line with the projections for moderate growth and subdued inflation in 2010 that they had submitted just before the Committee's November 3-4 meeting; accordingly, their views on the economic outlook had not changed appreciably. Participants expected the economic recovery to continue, but, consistent with experience following previous financial crises, most anticipated that the pickup in output and employment growth would be rather slow relative to past recoveries from deep recessions. A moderate pace of expansion would imply slow improvement in the labor market next year, with unemployment declining only gradually. Participants agreed that underlying inflation currently was subdued and was likely to remain so for some time. Some noted the risk that, over the next couple of years, inflation could edge further below the rates they judged most consistent with the Federal Reserve's dual mandate for maximum employment and price stability; others saw inflation risks as tilted toward the upside in the medium term.
21	0.46	2018-12-19	After taking into account incoming economic data, information from business contacts, and the tightening of financial conditions, participants generally revised down their individual assessments of the appropriate path for monetary policy and indicated either no material change or only a modest downward revision in their assessment of the economic outlook. Economic growth was expected to remain above trend in 2019 and then slow to a pace closer to trend over the medium term. Participants who downgraded their assessment of the economic outlook pointed to a variety of factors underlying their assessment, including recent financial market developments, some softening in the foreign economic growth outlook, or a more pessimistic outlook for housing-sector activity.

(continued)

Topic	Gamma	Date	Paragraph
21	0.68	2011-12-13	Members viewed the information on U.S. economic activity received over the intermeeting period as suggesting that the economy was expanding moderately. While overall labor market conditions had improved some in recent months, the unemployment rate remained elevated relative to levels that the Committee anticipated would prevail in the longer run. Inflation had moderated, and longer-term inflation expectations remained stable. However, available indicators pointed to some slowing in the pace of economic growth in Europe and in some emerging market economies. Members continued to expect a moderate pace of economic growth over coming quarters, with the unemployment rate declining only gradually toward levels consistent with the Committee's dual mandate. Strains in global financial markets continued to pose significant downside risks to economic activity. Members also anticipated that inflation would settle, over coming quarters, at levels at or below those consistent with the dual mandate.

Note: This table presents paragraphs categorized under the *Mandate* aggregate, identified using an STM with 21 topics. An overview of all 21 topics is presented in Table 1. The gamma value represents the probability that a paragraph belongs to a specific topic. Paragraphs are sorted by the topic with the highest gamma value.

A.2. Illustrative Case Study: Topic Granularity in FOMC Minutes

In our baseline analysis, we estimate an STM with $K = 21$ topics and subsequently aggregate selected topics into the broader categories *Mandate* and *Financial Stability*. One may ask why we do not directly estimate a substantially smaller number of topics that correspond more closely to these aggregate concepts. While such an approach would be more parsimonious, it risks combining paragraphs that are semantically related but economically distinct. In this section, we illustrate why a more granular topic specification is better suited to the structure of FOMC minutes and to the objective of our analysis.

For this purpose, we estimate an alternative STM with $K = 5$. Table A3 reports the share of paragraphs assigned to each topic and the corresponding highest-probability word stems. The five-topic model produces broad and semantically coherent themes. Topic 1 primarily captures monetary policy decisions and procedural language, while Topic 2 combines discussions of inflation, labor markets, and economic developments. Topic 3 also contains monetary policy, inflation, and economic-outlook terminology, indicating substantial conceptual overlap with Topics 1 and 2. Topic 4 combines business activity, consumption, investment, and housing, whereas Topic 5 contains terms related to financial markets, risk, credit, employment, and households.

Table A3: STM with 5 Topics ($K = 5$)

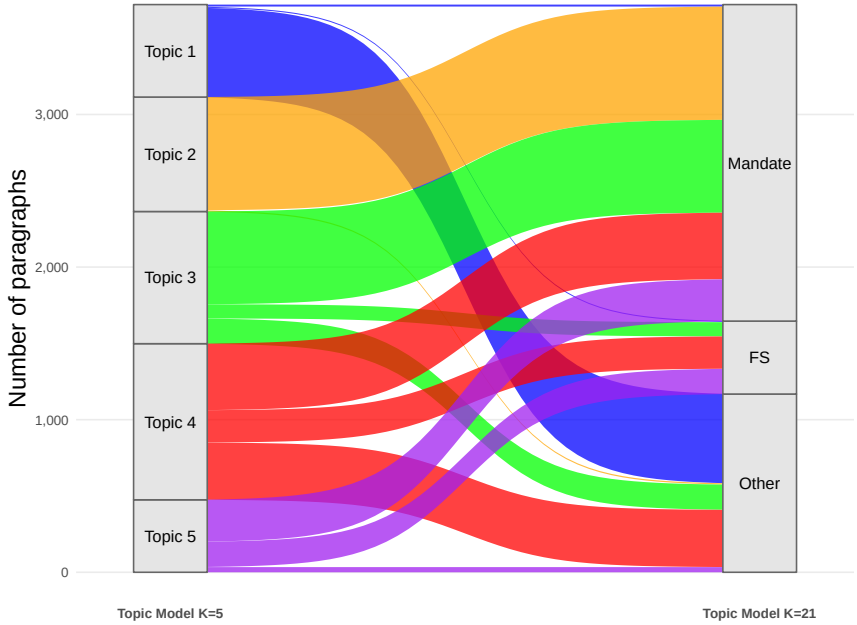
Topic	**Topic Prop.**	**Top Terms**
Topic 1	0.160	direct, committe, reserv, polici, growth, feder, rang, project, bank, meet, discuss, rate, vote, domest, conclus
Topic 2	0.205	inflat, price, labor, expect, market, recent, remain, continu, increas, particip, econom, pressur, growth, measur, term
Topic 3	0.223	polici, committe, rate, econom, feder, inflat, fund, monetari, particip, condit, agre, outlook, time, rang, market
Topic 4	0.262	busi, spend, growth, consum, activ, continu, hous, report, sector, recent, invest, demand, economi, increas, econom
Topic 5	0.150	particip, financi, market, econom, note, risk, condit, rate, credit, employ, continu, remain, activ, household, u.

Note: The table reports the 15 most probable word stems for each topic identified by the Structural Topic Model with $K = 5$. Topic proportions denote the share of paragraphs assigned to each topic.

The key question is whether these broad topics can replace the more granular topics used to construct our aggregate measures. Figure A2 maps the paragraphs assigned by the $K = 5$ model into the categories *Mandate*, *Financial Stability*, and *Other* obtained from the baseline $K = 21$ specification. The figure shows that each topic identified by the $K = 5$ model contributes to more than one of these categories. Conversely, neither *Mandate* nor *Financial Stability* can be recovered from a single low-dimensional topic. In particular, paragraphs classified as financial stability-related in the baseline model are drawn from

several of the five broad topics rather than from one clearly identifiable financial stability topic.

Figure A2: Comparison of Aggregated Topic Classifications in the 5 Topic and 21 Topic Models



Note: The figure compares the five topics obtained from the 5 topic model with the aggregated categories Mandate, Financial Stability (FS), and Other derived from the 21 topic model. Flow widths are proportional to the number of paragraphs. The flows show how paragraphs assigned to each topic in the 5 topic model are redistributed across the aggregated categories of the more granular 21 topic specification.

This exercise illustrates that high semantic coherence does not necessarily imply high economic interpretability. A low-dimensional topic model groups lexically related discussions into broad themes, but it does not sufficiently distinguish the economic concepts required for our analysis. Our two-step approach therefore first identifies more granular and differentiated topics and then aggregates economically related topics into theoretically motivated categories. For FOMC minutes, this approach provides a clearer separation between mandate-related, financial stability-related, and other discussions than directly imposing a small number of topics.

B. Robustness Checks and Further Extensions

In this appendix section, we present additional results, robustness checks, and methodological extensions. Our analysis combines machine learning-based textual measurement with econometrically estimated Taylor rules and sentiment regressions. Both components require a number of modeling choices concerning the construction of the textual indicators and the specification of the empirical models. Although these choices follow established practices in the literature, they inevitably involve some degree of researcher discretion. The following subsections therefore examine whether our main results and conclusions are sensitive to alternative measurement choices, model specifications, and empirical assumptions.

This section broadly follows the sequence of the main analysis. We begin by examining the construction of the textual variables. Specifically, we consider alternative aggregations of the STM topics (Section B.1), alternative numbers of topics in the STM estimation (Section B.2), separate positivity and negativity indicators instead of the baseline net-sentiment measure (Section B.3), and interactions between topic prevalence and topic-specific sentiment (Section B.4). Where applicable, we report the corresponding robustness results for both the Taylor rule estimation and sentiment regression analyses.

The remaining sections focus on alternative empirical specifications. Section B.5 examines the sensitivity of the results to alternative forecasting variables and horizons, financial stability indicators, shadow rate estimates, and Taylor rule specifications including two lags of the dependent variable. It also reports the corresponding robustness checks for the sentiment regressions where applicable. Finally, Section B.6 replaces the ZLB interaction specification with separate estimations across monetary policy regimes to assess whether the main Taylor rule results are driven by the imposed interaction terms or by particular episodes within the ZLB period.

B.1. Alternative Topic Aggregation for the Topic Proportion Variables

Table B1: Narrower Topic Aggregation for Mandate Discussions and Financial Stability

Procedural Information	**Financial Conditions**^{FS}	**Financial Vulnerabilities and Risk**^{FS}	**Business Investment and Production**	**Labor Market Conditions**^{M}	**Dual-Mandate Objectives**^{M}	**Household Sector**
releas	secur	financi	busi	labor	particip	consum
statement	feder	market	invest	particip	committe	spend
vote	reserv	particip	capit	market	polici	household
short	agenc	credit	report	rate	rate	incom
approv	hold	bank	spend	unemploy	inflat	recent
encompass	committe	risk	firm	employ	monetari	growth
p.m	treasuri	condit	contact	wage	goal	increas
meet	polici	note	equip	worker	feder	effect
inclus	agre	term	product	job	guidanc	support
press	term	increas	continu	increas	percent	confid
sentenc	support	loan	industri	condit	time	fiscal
paragraph	market	yield	construct	growth	run	sale
text	level	equiti	increas	forc	employ	gain
percentag	maintain	rate	profit	continu	maximum	factor
unanim	condit	remain	expenditur	remain	achiev	retail

Foreign Economic Activity	**Housing and Mortgages**	**Inflation Expectations**^{M}	**Voting Procedures**	**Real Activity and Demand**	**Policy Communication**	**Credit Easing**
u.	hous	inflat	direct	growth	committe	purchas
foreign	sector	expect	reserv	expans	statement	asset
export	particip	particip	feder	econom	meet	committe
economi	activ	term	polici	economi	polici	pace
econom	price	measur	committe	inventori	agre	program
growth	report	run	conclus	activ	econom	particip
domest	mortgag	percent	bank	demand	discuss	market
trade	home	market	domest	prospect	languag	econom
dollar	continu	committe	instruct	current	decis	outlook
demand	construct	month	author	develop	chang	progress
effect	contact	remain	accord	quarter	futur	addit
global	rate	continu	execut	moder	risk	condit
abroad	district	recent	transact	rate	postmeet	continu
develop	sale	price	york	outlook	announc	reduct
market	recent	object	discuss	busi	communic	improv

Money and Credit Aggregates	**Economic Risks and Shocks**	**Outlook Assessment**	**Policy Implementation**	**Policy Stance and Risks**	**Inflation Outlook**^{M}	**Growth Outlook and Risks**
rang	particip	project	rate	polici	inflat	econom
growth	econom	rate	fund	tighten	price	growth
committe	risk	gdp	feder	risk	increas	particip
reserv	activ	real	inflat	economi	pressur	risk
monetari	note	forecast	rang	committe	cost	outlook
econom	suppli	percent	committe	monetari	labor	pace
aggreg	economi	growth	target	action	product	continu
period	polici	run	condit	period	rise	expect
price	pandem	unemploy	outlook	economi	expect	meet
debt	continu	polici	polici	move	resourc	recent
condit	outlook	particip	expect	eas	energi	activ
consist	support	feder	percent	view	recent	downsid
accept	remain	assess	market	stanc	core	quarter
slight	fiscal	inflat	assess	intermeet	remain	moder
intermeet	uncertainti	monetari	direct	direct	market	anticip

Note: The table reports the 15 most influential word stems for each STM topic, ranked by the probabilities that the word appears in the respective topic. Topic labels are assigned based on dominant word stems and representative paragraphs (see Appendix A.1). Superscripts indicate topics included in the narrower aggregation used for the robustness analysis: M denotes *Mandate* and FS denotes *Financial Stability*.

Table B2: Taylor Rule Estimates with Narrow Topic Definition (2000 - 2018)

	FFR				Shadow Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Short-run estimates						
i_{t-1}	0.947***	0.949***	0.952***	0.935***	0.959***	0.964***
	(0.022)	(0.024)	(0.025)	(0.033)	(0.028)	(0.027)
$\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$	0.069	0.095	0.124	0.229*	0.116	0.093
	(0.085)	(0.090)	(0.093)	(0.121)	(0.111)	(0.111)
$\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$	0.034***	0.016	0.013	0.045*	0.074***	0.072***
	(0.012)	(0.011)	(0.014)	(0.024)	(0.026)	(0.027)
TP_t^M		0.266	0.255	0.368	0.298	0.226
		(0.192)	(0.185)	(0.280)	(0.292)	(0.330)
TP_t^{FS}		-0.740	-0.619	-2.002***	-1.792***	-1.702**
		(0.561)	(0.499)	(0.738)	(0.691)	(0.705)
$Sentiment_t^M$			0.024	0.090	0.093	0.093
			(0.105)	(0.111)	(0.112)	(0.112)
$Sentiment_t^{FS}$			0.173	0.139	0.158	0.144
			(0.125)	(0.182)	(0.172)	(0.172)
VIX_t						-0.003
						(0.008)
D_t^{ZLB}				-0.031	0.018	-0.017
				(0.180)	(0.234)	(0.241)
$D_t^{ZLB} \times TP_t^M$				-0.472	-0.396	-0.290
				(0.355)	(0.433)	(0.445)
$D_t^{ZLB} \times TP_t^{FS}$				2.911***	2.977***	3.025***
				(0.971)	(1.042)	(1.037)
$D_t^{ZLB} \times Sentiment_t^M$				-0.256	-0.185	-0.188
				(0.160)	(0.184)	(0.194)
$D_t^{ZLB} \times Sentiment_t^{FS}$				-0.077	0.009	0.013
				(0.201)	(0.211)	(0.214)
Constant	0.023	-0.059	-0.113	-0.206	-0.061	0.039
	(0.134)	(0.140)	(0.144)	(0.159)	(0.159)	(0.254)
Wald Tests (χ^2)						
$\theta_{TP^M} + \varphi_{TP^M} = 0$				0.198	0.074	0.030
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$				4.954**	3.756*	4.058**
$\theta_{Sent^M} + \varphi_{Sent^M} = 0$				1.720	0.331	0.294
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$				0.567	2.048	1.700
Sample	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018
N	152	152	152	152	152	152
$Adj.R^2$	0.971	0.971	0.971	0.975	0.982	0.982

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018 using the narrow topic definitions of mandate and financial stability discussions specified in Table B1. Columns (1)–(4) use the effective federal funds rate as the dependent variable. Columns (5)–(6) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period, while column (6) additionally controls for the VIX. $\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate-related and financial-stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment measures constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. The Wald test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table B3: Sentiment Regression with a Narrow Topic Definition (2011 - 2018)

	Statement			Q&A Session		
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.111	0.118	0.163	0.024	0.063	-0.008
	(0.113)	(0.147)	(0.155)	(0.052)	(0.094)	(0.088)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.018**	0.006	0.016	0.006	0.005	-0.012
	(0.008)	(0.014)	(0.016)	(0.005)	(0.007)	(0.014)
TP_t^M		-0.050	-0.096		0.023	0.096
		(0.160)	(0.161)		(0.133)	(0.149)
TP_t^{FS}		-0.123	-0.001		0.007	-0.187
		(0.315)	(0.332)		(0.340)	(0.329)
$Sentiment_t^M$		0.194**	0.164*		-0.013	0.034
		(0.082)	(0.087)		(0.078)	(0.084)
$Sentiment_t^{FS}$		0.011	0.004		0.039	0.050
		(0.055)	(0.056)		(0.054)	(0.051)
i_{t-1}			-0.022			0.035*
			(0.019)			(0.018)
Constant	-0.062	-0.054	-0.122	-0.059	-0.130	-0.022
	(0.194)	(0.280)	(0.290)	(0.087)	(0.165)	(0.151)
Sample	2011-2018	2011-2018	2011-2018	2011-2018	2011-2018	2011-2018
N	32	32	32	32	32	32
$Adj.R^2$	0.344	0.379	0.393	0.040	-0.073	0.158

Note: The table reports OLS estimates of equation (9) for the period from April 2011 to December 2018, using the alternative (narrow) topic definitions for mandate objectives and financial stability discussions reported in Table B1. Each observation corresponds to one press conference. The dependent variable is the sentiment of the Chair's introductory statement in columns (1)–(3) and the sentiment of the Chair's responses during the Q&A session in columns (4)–(6). $\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. TP_t^M and TP_t^{FS} are the proportions of FOMC discussions devoted to mandate objectives and financial stability, while $Sentiment_t^M$ and $Sentiment_t^{FS}$ are the corresponding topic-specific sentiment measures. i_{t-1} denotes the policy rate at the previous FOMC meeting, measured by the effective federal funds rate and replaced by the shadow rate of Wu and Xia (2016) during the ZLB period. Sentiment is measured using the fine-tuned RoBERTa model described in Section 3.3. Bootstrapped standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table B4: Topic Aggregation for Economic Outlook, Monetary Policy Inclination and Financial Stability

Procedural Information	**Financial Conditions**^{FS}	**Financial Vulnerabilities and Risk**^{FS}	**Business Investment and Production**	**Labor Market Conditions**	**Dual-Mandate Objectives**^{MPI}	**Household Sector**^{EO}
releas	secur	financi	busi	labor	particip	consum
statement	feder	market	invest	particip	committe	spend
vote	reserv	particip	capit	market	polici	household
short	agenc	credit	report	rate	rate	incom
approv	hold	bank	spend	unemploy	inflat	recent
encompass	committe	risk	firm	employ	monetari	growth
p.m	treasuri	condit	contact	wage	goal	increas
meet	polici	note	equip	worker	feder	effect
inclus	agre	term	product	job	guidanc	support
press	term	increas	continu	increas	percent	confid
sentenc	support	loan	industri	condit	time	fiscal
paragraph	market	yield	construct	growth	run	sale
text	level	equiti	increas	forc	employ	gain
percentag	maintain	rate	profit	continu	maximum	factor
unanim	condit	remain	expenditur	remain	achiev	retail

Foreign Economic Activity	**Housing and Mortgages**^{FS}	**Inflation Expectations**^{MPI}	**Voting Procedures**	**Real Activity and Demand**^{EO}	**Policy Communication**	**Credit Easing**^{MPI}
u.	hous	inflat	direct	growth	committe	purchas
foreign	sector	expect	reserv	expans	statement	asset
export	particip	particip	feder	econom	meet	committe
economi	activ	term	polici	economi	polici	pace
econom	price	measur	committe	inventori	agre	program
growth	report	run	conclus	activ	econom	particip
domest	mortgag	percent	bank	demand	discuss	market
trade	home	market	domest	prospect	languag	econom
dollar	continu	committe	instruct	current	decis	outlook
demand	construct	month	author	develop	chang	progress
effect	contact	remain	accord	quarter	futur	addit
global	rate	continu	execut	moder	risk	condit
abroad	district	recent	transact	rate	postmeet	continu
develop	sale	price	york	outlook	announc	reduct
market	recent	object	discuss	busi	communic	improv

Money and Credit Aggregates	**Economic Risks and Shocks**^{EO}	**Outlook Assessment**	**Policy Implementation**^{MPI}	**Policy Stance and Risks**^{MPI}	**Inflation Outlook**^{MPI}	**Growth Outlook and Risks**^{EO}
rang	particip	project	rate	polici	inflat	econom
growth	econom	rate	fund	tighten	price	growth
committe	risk	gdp	feder	risk	increas	particip
reserv	activ	real	inflat	economi	pressur	risk
monetari	note	forecast	rang	committe	cost	outlook
econom	suppli	percent	committe	monetari	labor	pace
aggreg	economi	growth	target	action	product	continu
period	polici	run	condit	period	rise	expect
price	pandem	unemploy	outlook	economi	expect	meet
debt	continu	polici	polici	move	resourc	recent
condit	outlook	particip	expect	eas	energi	activ
consist	support	feder	percent	view	recent	downsid
accept	remain	assess	market	stanc	core	quarter
slight	fiscal	inflat	assess	intermeet	remain	moder
intermeet	uncertainti	monetari	direct	direct	market	anticip

Note: The table reports the 15 most influential word stems for each STM topic, ranked by the probabilities that the word appears in the respective topic. Topic labels are assigned based on dominant word stems and representative paragraphs (see Appendix A.1). Superscripts indicate topics included in the alternative aggregation used for the robustness analysis: *EO* denotes *Economic Outlook*, *MPI* denotes *Monetary Policy Inclination*, and *FS* denotes *Financial Stability*. The *Outlook Assessment* topic is not included in the Economic Outlook aggregation because it consists primarily of procedural paragraphs stating that participants submitted their individual projections for the SEP, without reporting their substantive content, alongside a smaller number of paragraphs concerning the preparation of the semiannual report to Congress.

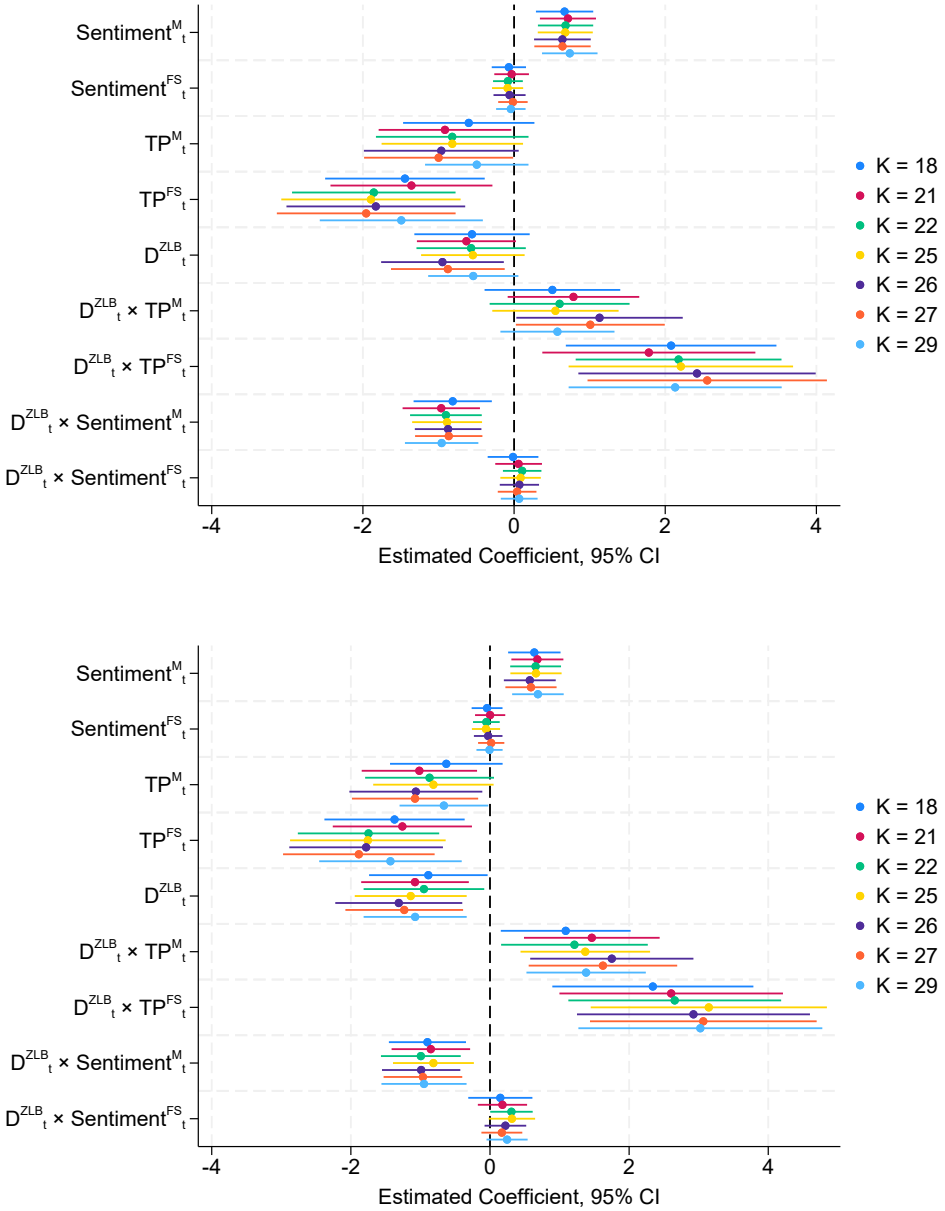
Table B5: Taylor Rule Estimates with Alternative Topic Aggregation (2000–2018)

	FFR	Shadow Rate	
	(1)	(2)	(3)
Short-run estimates			
i_{t-1}	0.963***	0.981***	0.983***
	(0.031)	(0.028)	(0.028)
$\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$	0.253*	0.170	0.153
	(0.130)	(0.122)	(0.123)
$\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$	0.060**	0.092***	0.091***
	(0.030)	(0.031)	(0.033)
TP_t^O	-0.845	-0.839	-0.794
	(0.558)	(0.547)	(0.605)
TP_t^I	-0.898*	-1.027**	-0.992*
	(0.499)	(0.492)	(0.523)
TP_t^{FS}	-1.355***	-1.243***	-1.196**
	(0.501)	(0.475)	(0.508)
$Sentiment_t^O$	0.221**	0.231**	0.226**
	(0.110)	(0.110)	(0.111)
$Sentiment_t^I$	0.283**	0.277*	0.275*
	(0.143)	(0.143)	(0.142)
$Sentiment_t^{FS}$	0.016	0.040	0.038
	(0.126)	(0.122)	(0.120)
VIX			-0.002
			(0.008)
D_t^{ZLB}	-0.274	-0.966*	-0.958*
	(0.468)	(0.537)	(0.560)
$D_t^{ZLB} \times TP_t^O$	0.929	1.946**	1.907**
	(0.770)	(0.816)	(0.893)
$D_t^{ZLB} \times TP_t^I$	0.260	1.310*	1.296*
	(0.610)	(0.709)	(0.750)
$D_t^{ZLB} \times TP_t^{FS}$	1.442**	2.472***	2.480***
	(0.704)	(0.796)	(0.808)
$D_t^{ZLB} \times Sentiment_t^O$	-0.253*	-0.149	-0.150
	(0.139)	(0.165)	(0.170)
$D_t^{ZLB} \times Sentiment_t^I$	-0.479**	-0.581**	-0.559**
	(0.227)	(0.255)	(0.258)
$D_t^{ZLB} \times Sentiment_t^{FS}$	-0.044	-0.002	-0.007
	(0.170)	(0.186)	(0.195)
$Constant$	0.350	0.481	0.505
	(0.401)	(0.402)	(0.396)
Wald Tests (χ^2)			
$\theta_{TP^O} + \varphi_{TP^O} = 0$	0.025	3.208*	3.110*
$\theta_{TP^I} + \varphi_{TP^I} = 0$	1.757	0.203	0.236
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$	0.036	4.022**	4.194**
$\theta_{Sent^O} + \varphi_{Sent^O} = 0$	0.187	0.533	0.400
$\theta_{Sent^I} + \varphi_{Sent^I} = 0$	1.406	2.130	1.604
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$	0.061	0.073	0.047
Sample	2000–2018	2000–2018	2000–2018
N	152	152	152
$Adj.R^2$	0.975	0.982	0.982

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018. Following the distinction proposed by Ehrmann and Fratzscher (2007), discussions related to the Federal Reserve's mandate are decomposed into *Economic Outlook* (O) and *Monetary Policy Inclination* (I). Column (1) uses the effective federal funds rate as the dependent variable. Columns (2)–(3) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period, while column (3) additionally controls for the VIX. $\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^O, TP_t^I, and TP_t^{FS} denote the topic proportions of economic outlook, monetary policy inclination, and financial-stability discussions, respectively. $Sentiment_t^O$, $Sentiment_t^I$, and $Sentiment_t^{FS}$ denote the corresponding sentiment measures, constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. The Wald test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

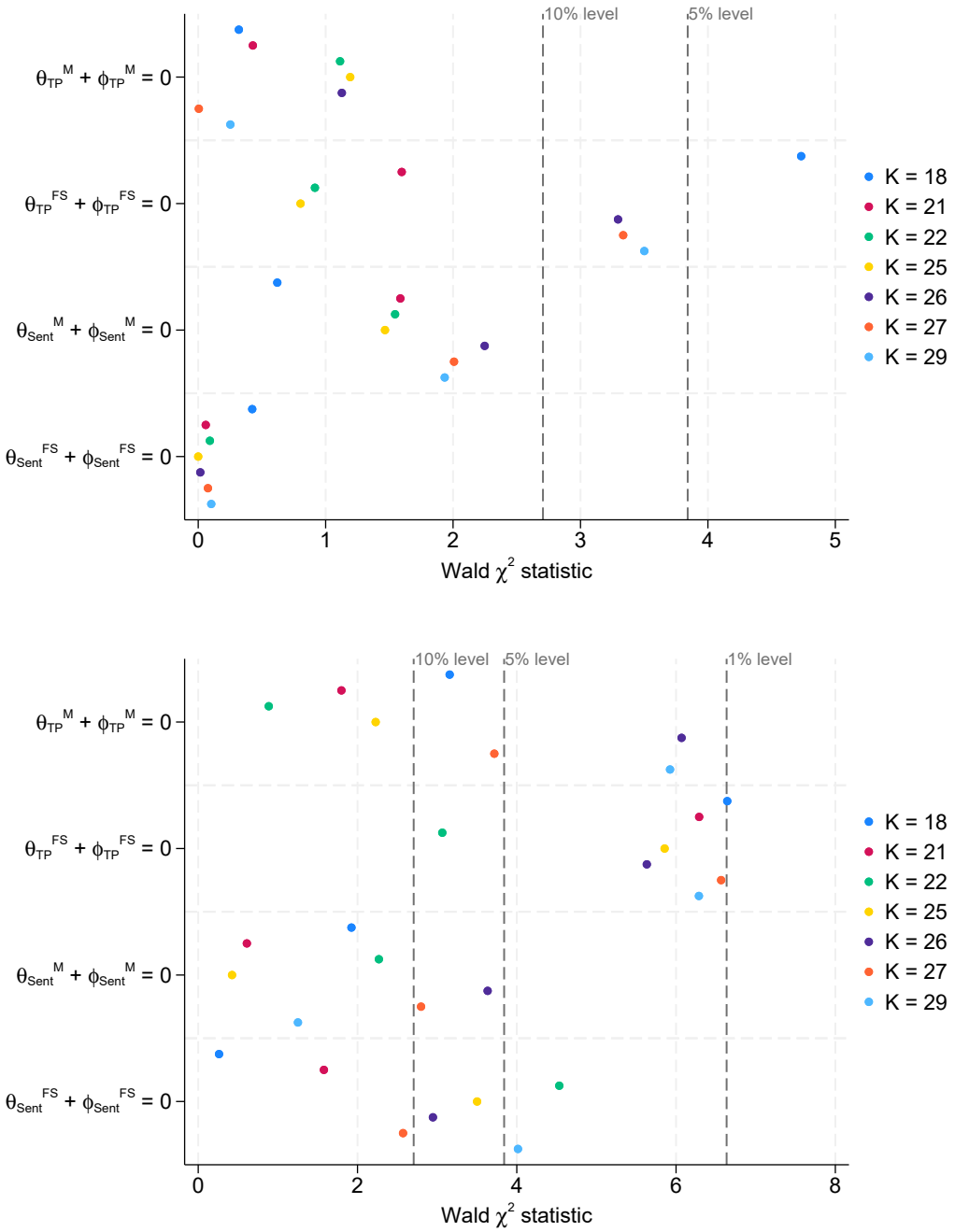
B.2. Alternative Numbers of Topics for STM Estimation

Figure B1: Coefficient Plots for Taylor-Rule Estimates Using Alternative Numbers of STM Topics



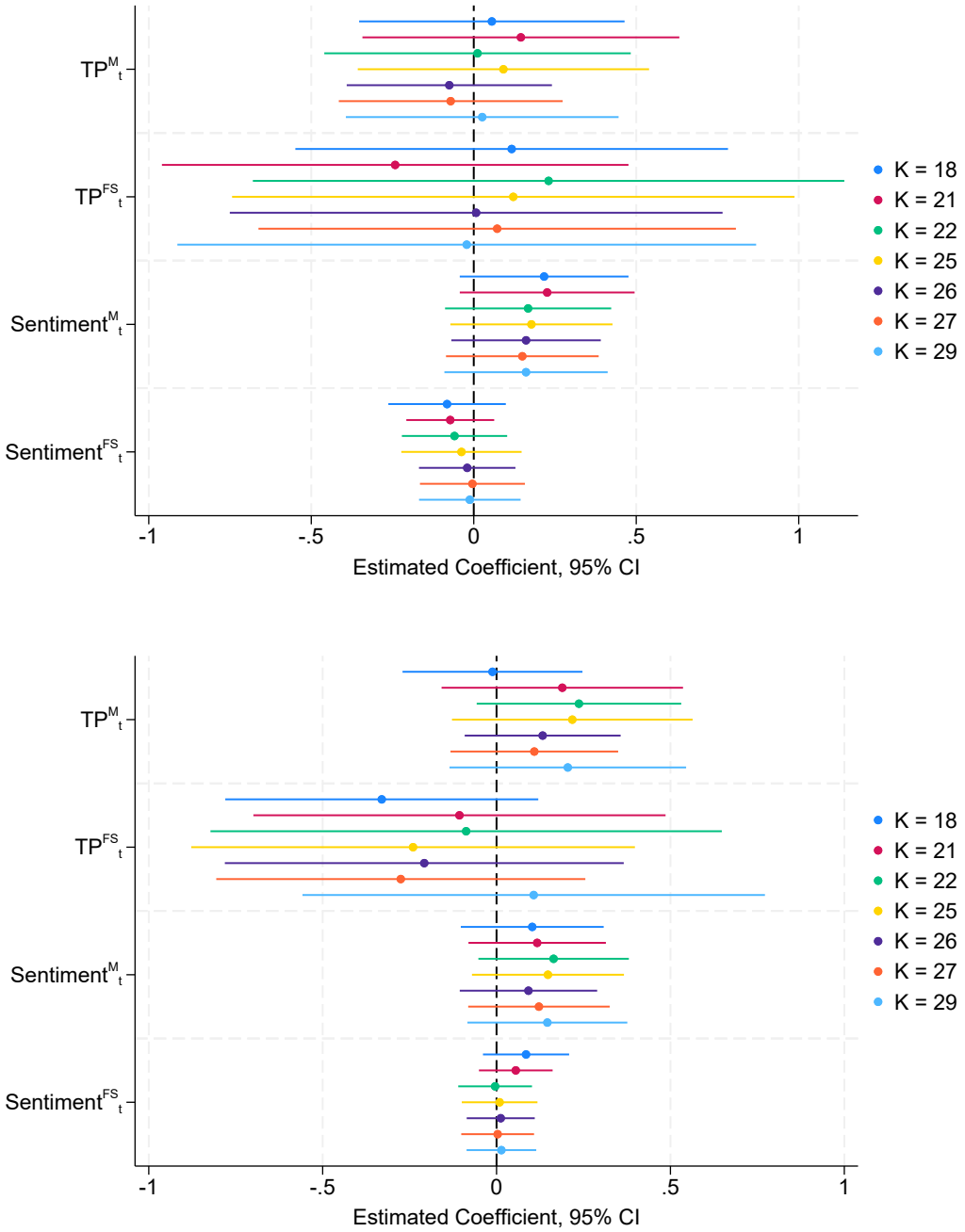
Note: The figure shows the regression coefficients for different topic selections $K \in \{18, 21, 22, 25, 26, 27, 29\}$ derived from the STM. The colored points represent the coefficient estimates, while the lines indicate the 95% confidence intervals. The coefficients correspond to the topic proportions (TP_t^M, TP_t^{FS}) and sentiment indicators ($Sentiment_t^M$, $Sentiment_t^{FS}$) associated with *Mandate* and *Financial Stability* topics, respectively. D_t^{ZLB} is a dummy variable for the ZLB period. The ZLB dummy is set to one for the period from December 2008 to December 2015. The specification in the first figure corresponds to the results shown in Table 4, column (4), while the specification in the second figure uses shadow rates for the ZLB period with the same specification (column (5)).

Figure B2: Wald Tests for Taylor-Rule Estimates Using Alternative Numbers of STM Topics



Note: The figures report Wald χ^2 statistics for the null hypotheses $H_0 : \theta_j + \varphi_j = 0$, where $j \in \{TP^M, TP^{FS}, Sentiment^M, Sentiment^{FS}\}$, across alternative STM topic selections $K \in \{18, 21, 22, 25, 26, 27, 29\}$. Here, θ_j denotes the baseline coefficient and φ_j its interaction with D_t^{ZLB}. Hence, $\theta_j + \varphi_j$ is the estimated effect during the ZLB period, defined as December 2008 to December 2015. Larger values on the horizontal axis indicate stronger evidence against the null hypothesis. The vertical lines denote the χ^2 critical values at the 10%, 5%, and 1% significance levels, respectively. The first figure corresponds to column (4) of Table 4, while the second uses the shadow rate during the ZLB period, as in column (5).

Figure B3: Coefficient Plots for Sentiment Regressions Using Alternative Numbers of STM Topics



Note: The figure reports coefficient estimates from the sentiment regressions using alternative numbers of topics, $K \in \{18, 21, 22, 25, 26, 27, 29\}$, in the STM. The points represent the estimated coefficients, and the horizontal lines indicate the corresponding 95% confidence intervals. The reported coefficients refer to the topic proportions (TP_t^M and TP_t^{FS}) and the topic-specific sentiment measures ($Sentiment_t^M$ and $Sentiment_t^{FS}$) for mandate-related and financial stability discussions. The upper panel reports the results for the sentiment of the Chair's introductory statement and corresponds to the specification in column (3) of Table 5. The lower panel reports the results for the sentiment of the Chair's responses during the Q&A session and corresponds to the specification in column (6). All remaining controls are identical to those in the respective baseline specifications.

B.3. Replacing the Net-Sentiment Indicator with Positivity and Negativity Indicators

In this robustness exercise, we replace the aggregate net-sentiment measure, which captures the difference between the shares of positive and negative sentences, with separate positivity and negativity indicators. This allows us to relax the symmetry assumption imposed by the baseline measure and to examine whether positive and negative sentiment are associated differently with monetary policy decisions. For each discussion category, the indicators are defined as the respective shares of positive and negative sentences among all classified sentences. Specifically, we calculate

$$Sentiment_t^{Positive,T} = \frac{N_t^{Positive,T}}{N_t^{Positive,T} + N_t^{Neutral,T} + N_t^{Negative,T}}, \qquad T \in \{\mathrm{M}, \mathrm{FS}\}, \tag{B.1}$$

and

$$Sentiment_t^{Negative,T} = \frac{N_t^{Negative,T}}{N_t^{Positive,T} + N_t^{Neutral,T} + N_t^{Negative,T}}, \qquad T \in \{\mathrm{M}, \mathrm{FS}\}, \tag{B.2}$$

where M denotes mandate-related discussions and FS denotes financial stability-related discussions.

Table B6 reports the Taylor rule estimates obtained when the baseline net-sentiment measure in Equation (6) is replaced by the negativity indicators defined in Equation (B.2).

Table B7 reports the results from a more flexible specification in which the positivity and negativity indicators defined in Equations (B.1) and (B.2) are included jointly. The first two Wald tests again evaluate the total effects of the mandate-related and financial stability-related topic proportions at the ZLB. We additionally test whether positive and negative sentiment have effects of equal magnitude and opposite sign. Outside the ZLB, the symmetry tests evaluate

$$H_0 : \theta_{Positive,T} + \theta_{Negative,T} = 0, \qquad T \in \{M, FS\}.$$

During the ZLB period, the corresponding tests account for both the baseline coefficients and their interactions with the ZLB indicator:

$$H_0 : (\theta_{Positive,T} + \varphi_{Positive,T}) + (\theta_{Negative,T} + \varphi_{Negative,T}) = 0, \qquad T \in \{M, FS\}.$$

All reported Wald-test values are χ^2 statistics.

Table B6: Taylor Rule Estimates with Negativity Indicators (2000–2018)

	FFR				Shadow Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Short-run Estimates						
i_{t-1}	0.947***	0.929***	0.951***	0.945***	0.971***	0.978***
	(0.022)	(0.025)	(0.024)	(0.031)	(0.026)	(0.026)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.069	0.133	0.158	0.295**	0.188*	0.140
	(0.085)	(0.095)	(0.104)	(0.115)	(0.109)	(0.111)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.034***	0.041***	0.025	0.049*	0.078***	0.072***
	(0.012)	(0.014)	(0.018)	(0.025)	(0.027)	(0.028)
TP_t^M		-0.755**	-0.748**	-1.089**	-1.174***	-1.131***
		(0.300)	(0.321)	(0.436)	(0.418)	(0.421)
TP_t^{FS}		-0.682*	-0.412	-1.700***	-1.532***	-1.407**
		(0.368)	(0.375)	(0.623)	(0.569)	(0.555)
$NegSentiment_t^M$			-0.573**	-1.151***	-1.093***	-1.012***
			(0.292)	(0.355)	(0.362)	(0.348)
$NegSentiment_t^{FS}$			-0.057	0.103	0.040	0.036
			(0.171)	(0.201)	(0.189)	(0.189)
VIX_t						-0.005
						(0.007)
D_t^{ZLB}				-1.426***	-1.672***	-1.678***
				(0.401)	(0.461)	(0.465)
$D_t^{ZLB} \times TP_t^M$				1.050**	1.647***	1.665***
				(0.459)	(0.517)	(0.519)
$D_t^{ZLB} \times TP_t^{FS}$				1.937***	2.715***	2.747***
				(0.751)	(0.862)	(0.859)
$D_t^{ZLB} \times NegSentiment_t^M$				1.740***	1.412**	1.346**
				(0.472)	(0.556)	(0.530)
$D_t^{ZLB} \times NegSentiment_t^{FS}$				-0.049	-0.133	-0.118
				(0.261)	(0.269)	(0.273)
Constant	0.023	0.526**	0.635**	0.934***	1.090***	1.180***
	(0.134)	(0.249)	(0.276)	(0.354)	(0.360)	(0.384)
Wald Tests (χ^2)						
$\theta_{TP^M} + \varphi_{TP^M} = 0$				0.040	2.089	2.591
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$				0.550	4.121**	5.005**
$\theta_{Sentiment^{Negative,M}} + \varphi_{Sentiment^{Negative,M}} = 0$				2.974*	0.500	0.443
$\theta_{Sentiment^{Negative,FS}} + \varphi_{Sentiment^{Negative,FS}} = 0$				0.090	0.199	0.153
Sample	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018
N	152	152	152	152	152	152
Adj.R^2	0.971	0.972	0.973	0.977	0.983	0.983

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018. Columns (1)–(4) use the effective federal funds rate as the dependent variable. Columns (5)–(6) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period, while column (6) additionally controls for the VIX. $\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate and financial stability discussions, respectively. In this specification, the baseline net sentiment measure is replaced by $NegSentiment_t^T$, $T \in \{M, FS\}$, which denotes the share of negative sentences within the corresponding discussion category, as defined in Equation (B.2). The negativity indicators are constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. The Wald test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table B7: Taylor Rule Estimates with Separate Positive and Negative Sentiment Indicators (2000–2018)

	FFR				Shadow Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Short-run Estimates						
i_{t-1}	0.947***	0.929***	0.942***	0.939***	0.965***	0.970***
	(0.022)	(0.025)	(0.025)	(0.032)	(0.027)	(0.027)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.069	0.133	0.161	0.312**	0.210*	0.177
	(0.085)	(0.095)	(0.110)	(0.126)	(0.114)	(0.113)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.034***	0.041***	0.027	0.053**	0.081***	0.077***
	(0.012)	(0.014)	(0.018)	(0.026)	(0.028)	(0.029)
TP_t^M		-0.755**	-0.580*	-0.882*	-0.962**	-0.941**
		(0.300)	(0.323)	(0.468)	(0.452)	(0.458)
TP_t^{FS}		-0.682*	-0.359	-1.523**	-1.359**	-1.282**
		(0.368)	(0.379)	(0.640)	(0.585)	(0.584)
$PosSentiment_t^M$			0.723**	0.930	0.939	0.888
			(0.355)	(0.635)	(0.621)	(0.602)
$PosSentiment_t^{FS}$			-0.047	0.073	0.085	0.081
			(0.139)	(0.165)	(0.161)	(0.158)
$NegSentiment_t^M$			-0.174	-0.555	-0.493	-0.473
			(0.296)	(0.487)	(0.493)	(0.497)
$NegSentiment_t^{FS}$			0.022	0.162	0.099	0.093
			(0.179)	(0.211)	(0.198)	(0.197)
VIX_t						-0.004
						(0.007)
D_t^{ZLB}				-0.831	-1.097	-1.058
				(0.669)	(0.846)	(0.877)
$D_t^{ZLB} \times TP_t^M$				0.877*	1.439**	1.441**
				(0.501)	(0.592)	(0.600)
$D_t^{ZLB} \times TP_t^{FS}$				1.807**	2.750***	2.761***
				(0.805)	(0.891)	(0.895)
$D_t^{ZLB} \times PosSentiment_t^M$				-0.784	-0.989	-1.011
				(0.718)	(0.841)	(0.877)
$D_t^{ZLB} \times PosSentiment_t^{FS}$				0.026	0.307	0.255
				(0.226)	(0.327)	(0.355)
$D_t^{ZLB} \times NegSentiment_t^M$				1.246**	0.775	0.729
				(0.631)	(0.754)	(0.774)
$D_t^{ZLB} \times NegSentiment_t^{FS}$				-0.052	-0.073	-0.087
				(0.289)	(0.313)	(0.315)
$Constant$	0.023	0.526**	0.112	0.155	0.292	0.393
	(0.134)	(0.249)	(0.378)	(0.672)	(0.656)	(0.643)
Wald tests (χ^2)						
$\theta_{TP^M} + \varphi_{TP^M} = 0$				0.000	1.655	1.720
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$				0.574	5.142**	5.429**
Sentiment symmetry outside ZLB: M			1.008	0.129	0.187	0.165
Sentiment symmetry outside ZLB: FS			0.010	0.655	0.436	0.408
Sentiment symmetry at ZLB: M				2.302	0.043	0.012
Sentiment symmetry at ZLB: FS				0.484	0.812	0.516
Sample	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018
N	152	152	152	152	152	152
Adjusted R^2	0.971	0.972	0.973	0.977	0.983	0.983

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018. Columns (1)–(4) use the effective federal funds rate as the dependent variable. Columns (5)–(6) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period, while column (6) additionally controls for the VIX. $\mathbb{E}_t(\bar{\pi}_t^{1:4})$ and $\mathbb{E}_t(\bar{y}_t^{1:4})$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate and financial stability discussions, respectively. In this specification, the baseline net sentiment measure is replaced by the jointly included $PosSentiment_t^T$ and $NegSentiment_t^T$, $T \in \{M, FS\}$, which denote the shares of positive and negative sentences, respectively, within the corresponding discussion category, as defined in Equations (B.1) and (B.2). The positive and negative sentiment indicators are constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. The Wald test rows report χ^2 statistics for the reported linear restrictions on topic effects and sentiment symmetry, both outside and during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

B.4. Interaction of Topic Prevalence and Sentiment

To examine whether the effect of sentiment depends on the prevalence of the corresponding topic, we extend the baseline Taylor rule specification in Equation (6). Specifically, we interact each sentiment indicator with the topic proportion of the corresponding discussion category. We further interact these terms with the ZLB indicator, allowing the conditional relationship between topic prevalence and sentiment to differ across policy regimes. The extended specification retains all constituent lower-order terms and is given by

$$
\begin{aligned}
i_t = \alpha &+ \rho i_{t-1} + \beta^{SR}\mathbb{E}_t\big(\bar{\pi}_t^{1:4}\big) + \gamma^{SR}\mathbb{E}_t\big(\bar{y}_t^{1:4}\big) \\
&+ \theta_{TP^M} TP_t^M + \theta_{TP^{FS}} TP_t^{FS} + \theta_{Sent^M} Sentiment_t^M + \theta_{Sent^{FS}} Sentiment_t^{FS} \\
&+ \kappa_M \left(TP_t^M \times Sentiment_t^M\right) + \kappa_{FS}\left(TP_t^{FS} \times Sentiment_t^{FS}\right) \\
&+ \varphi_{ZLB} D_t^{ZLB} \\
&+ D_t^{ZLB}\Big[\varphi_{TP^M} TP_t^M + \varphi_{TP^{FS}} TP_t^{FS} + \varphi_{Sent^M} Sentiment_t^M + \varphi_{Sent^{FS}} Sentiment_t^{FS}\Big] \\
&+ D_t^{ZLB}\Big[\psi_M\left(TP_t^M \times Sentiment_t^M\right) + \psi_{FS}\left(TP_t^{FS} \times Sentiment_t^{FS}\right)\Big] + u_t.
\end{aligned}
\tag{B.3}
$$

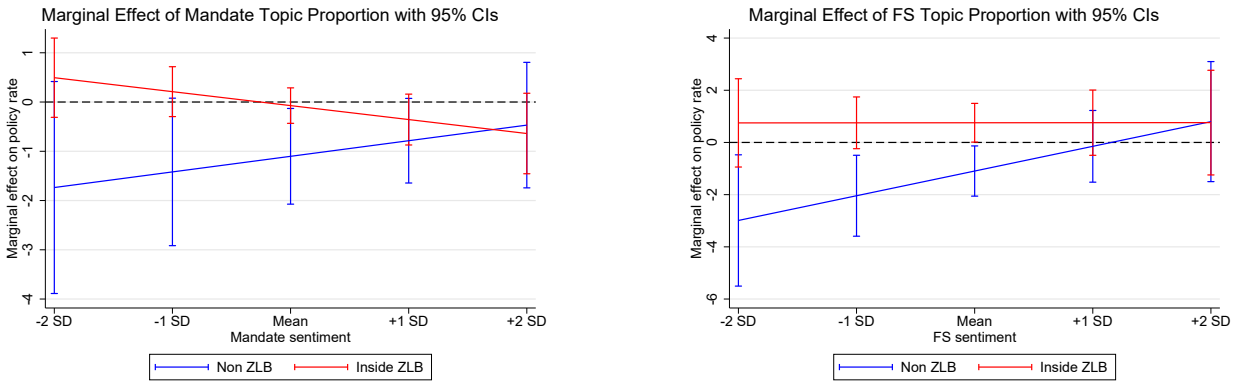
The coefficients κ_M and κ_{FS} capture the interaction between topic prevalence and sentiment outside the ZLB. At the ZLB, the corresponding interaction effects are given by $\kappa_M + \psi_M$ and $\kappa_{FS} + \psi_{FS}$, respectively. To illustrate the marginal effects of topic proportions, conditional on sentiment and the ZLB, we present margin plots in Figure B4.

Table B8: Taylor Rule Estimates with Topic Prevalence and Sentiment Interactions (2000-2018)

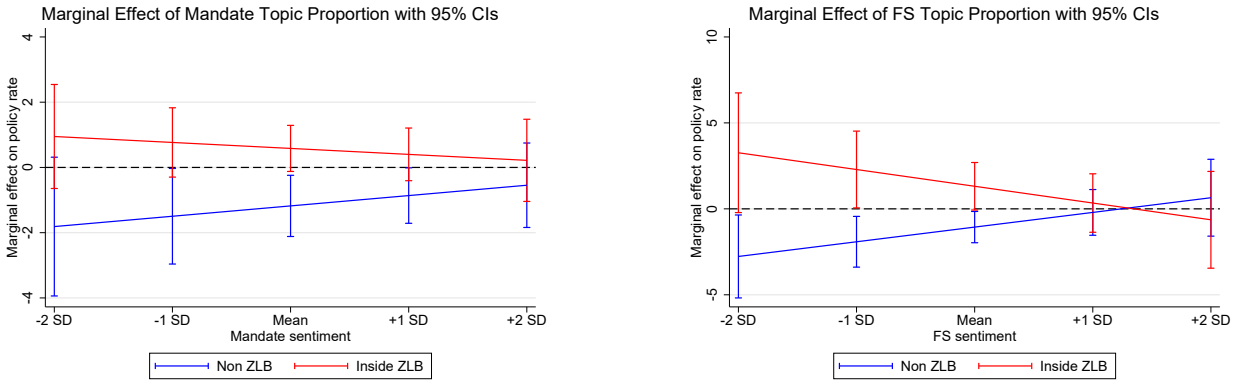
	FFR				Shadow Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Short-run Estimates:						
i_{t-1}	0.947***	0.929***	0.956***	0.949***	0.969***	0.976***
	(0.022)	(0.025)	(0.022)	(0.026)	(0.024)	(0.024)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.069	0.133	0.197*	0.288**	0.208*	0.161
	(0.085)	(0.095)	(0.105)	(0.118)	(0.114)	(0.112)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.034***	0.041***	0.012	0.046*	0.071***	0.066**
	(0.012)	(0.014)	(0.018)	(0.025)	(0.027)	(0.029)
TP_t^M		-0.755**	-0.706**	-1.102**	-1.179**	-1.144**
		(0.300)	(0.320)	(0.496)	(0.478)	(0.489)
TP_t^{FS}		-0.682*	-0.432	-1.095**	-1.061**	-0.981**
		(0.368)	(0.317)	(0.491)	(0.464)	(0.483)
$Sentiment_t^M$			0.545***	0.650***	0.633***	0.587***
			(0.165)	(0.176)	(0.180)	(0.170)
$Sentiment_t^{FS}$			-0.038	0.012	0.039	0.037
			(0.099)	(0.118)	(0.112)	(0.110)
$TP_t^M \times Sentiment_t^M$			1.907**	1.480	1.480	1.486
			(0.780)	(1.761)	(1.776)	(1.768)
$TP_t^{FS} \times Sentiment_t^{FS}$			3.048*	2.828*	2.551	2.424
			(1.574)	(1.684)	(1.631)	(1.584)
VIX_t						-0.005
						(0.007)
D_t^{ZLB}				0.131	0.199	0.190
				(0.092)	(0.130)	(0.133)
$D_t^{ZLB} \times TP_t^M$				1.030**	1.762***	1.800***
				(0.501)	(0.580)	(0.588)
$D_t^{ZLB} \times TP_t^{FS}$				1.850***	2.374**	2.324**
				(0.714)	(0.960)	(0.966)
$D_t^{ZLB} \times Sentiment_t^M$				-1.097***	-0.878***	-0.846***
				(0.247)	(0.286)	(0.291)
$D_t^{ZLB} \times Sentiment_t^{FS}$				0.049	0.318*	0.329*
				(0.168)	(0.186)	(0.184)
$D_t^{ZLB} \times TP_t^M \times Sentiment_t^M$				-2.803	-2.333	-2.274
				(1.934)	(2.246)	(2.286)
$D_t^{ZLB} \times TP_t^{FS} \times Sentiment_t^{FS}$				-2.820	-5.464**	-6.138**
				(2.152)	(2.615)	(2.971)
$Constant$	0.023	-0.041	-0.232	-0.383**	-0.296	-0.136
	(0.134)	(0.143)	(0.178)	(0.188)	(0.184)	(0.235)
Wald Tests (χ^2)						
$\theta_{TP^M} + \varphi_{TP^M} = 0$				0.156	2.620	3.083*
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$				3.980**	3.465*	3.572*
$\theta_{Sent^M} + \varphi_{Sent^M} = 0$				3.755*	0.877	0.905
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$				0.228	5.673**	5.968**
Sample	2000–2018	2000–2018	2000–2018	2000–2018	2000–2018	2000–2018
N	152	152	152	152	152	152
$Adj.R^2$	0.971	0.972	0.975	0.978	0.984	0.984

Note: The table reports OLS estimates of the baseline and extended specifications in equation (6) and equation (B.3) for the period from January 2000 to December 2018. Columns (1)–(4) use the effective federal funds rate as the dependent variable. Columns (5)–(6) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period, while column (6) additionally controls for the VIX. $\mathbb{E}_t(\bar{\pi}_t^{1:4})$ and $\mathbb{E}_t(\bar{y}_t^{1:4})$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate and financial stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment measures, constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. The interaction terms $TP_t^M \times Sentiment_t^M$ and $TP_t^{FS} \times Sentiment_t^{FS}$ capture whether the effect of sentiment depends on the prevalence of the corresponding discussion topic. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. The Wald test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Margin Plot for Column (4) of Table B8



Margin Plot for Column (5) of Table B8



Margin Plot for Column (6) of Table B8

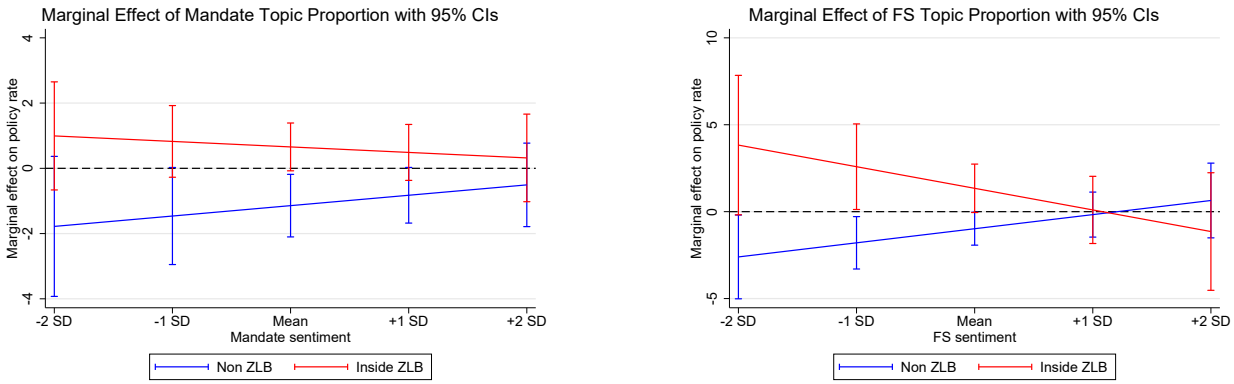


Figure B4: These figures present the marginal effects of topic proportions, conditional on sentiment, from the specifications in Columns (4)–(6) of Table B8. In each row, the left panel shows the marginal effect of increasing the *Mandate* topic proportion conditional on sentiment and ZLB status, while the right panel shows the marginal effect of increasing the *Financial Stability* topic proportion conditional on financial-stability sentiment and ZLB status. Error bars indicate 95% confidence intervals. The ZLB indicator equals one from December 2008 through December 2015 and zero otherwise.

B.5. Alternative Empirical Specifications

Table B9: Taylor Rule Estimates with Unemployment Gap Forecasts (2000–2018)

	FFR				Shadow Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Long-run estimates						
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	1.684	2.315**	4.014*	6.370**	7.963	8.240
	(1.471)	(1.075)	(2.124)	(2.712)	(6.121)	(8.431)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.485	0.395*	0.223	0.354	1.647	2.038
	(0.298)	(0.216)	(0.338)	(0.393)	(1.875)	(2.941)
Constant	-0.337	6.152	4.678	3.114	12.298	23.489
	(2.730)	(3.988)	(6.654)	(6.340)	(15.809)	(32.284)
Short-run estimates						
i_{t-1}	0.950***	0.932***	0.953***	0.944***	0.971***	0.979***
	(0.022)	(0.025)	(0.022)	(0.028)	(0.026)	(0.026)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.085	0.157	0.190	0.359***	0.230*	0.177
	(0.090)	(0.099)	(0.119)	(0.127)	(0.126)	(0.122)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.024**	0.027**	0.011	0.020	0.048**	0.044**
	(0.010)	(0.012)	(0.015)	(0.018)	(0.020)	(0.021)
TP_t^M		-0.674**	-0.575*	-0.754*	-0.818**	-0.798**
		(0.292)	(0.317)	(0.429)	(0.402)	(0.406)
TP_t^{FS}		-0.698*	-0.357	-1.287**	-1.121**	-1.021**
		(0.374)	(0.349)	(0.545)	(0.514)	(0.504)
$Sentiment_t^M$			0.485***	0.788***	0.767***	0.707***
			(0.171)	(0.195)	(0.198)	(0.177)
$Sentiment_t^{FS}$			-0.021	-0.041	-0.013	-0.011
			(0.105)	(0.123)	(0.118)	(0.118)
VIX_t						-0.006
						(0.007)
D_t^{ZLB}				-0.640*	-1.037***	-1.071***
				(0.332)	(0.396)	(0.400)
$D_t^{ZLB} \times TP_t^M$				0.701*	1.307***	1.345***
				(0.424)	(0.482)	(0.484)
$D_t^{ZLB} \times TP_t^{FS}$				1.716**	2.454***	2.481***
				(0.693)	(0.811)	(0.796)
$D_t^{ZLB} \times Sentiment_t^M$				-0.950***	-0.847***	-0.827***
				(0.253)	(0.287)	(0.274)
$D_t^{ZLB} \times Sentiment_t^{FS}$				0.076	0.189	0.155
				(0.155)	(0.184)	(0.198)
Constant	-0.017	0.417*	0.221	0.176	0.355	0.505
	(0.141)	(0.242)	(0.310)	(0.371)	(0.366)	(0.385)
Wald Tests (χ^2)						
$\theta_{TP^M} + \varphi_{TP^M} = 0$				0.083	2.224	2.703
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$				2.088	6.465**	7.189***
$\theta_{Sent^M} + \varphi_{Sent^M} = 0$				0.883	0.155	0.334
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$				0.112	1.454	0.876
Sample	2000–2018	2000–2018	2000–2018	2000–2018	2000–2018	2000–2018
N	152	152	152	152	152	152
Adj. R^2	0.970	0.971	0.973	0.976	0.983	0.983

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018. Columns (1)–(4) use the effective federal funds rate as the dependent variable. Columns (5)–(6) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period, while column (6) additionally controls for the VIX. $\mathbb{E}_t(\bar{\pi}_t^{1:4})$ denotes the average of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation, while $\mathbb{E}_t(\bar{y}_t^{1:4})$ denotes the average of the one- through four-quarter-ahead Greenbook forecasts of the unemployment gap. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate and financial stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment measures, constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. Long-run estimates are obtained by dividing the corresponding short-run coefficients by $1-\rho$, where ρ denotes the estimated coefficient on the lagged dependent variable. The Wald test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table B10: Sentiment Regression with Unemployment Gap Forecast (2011 - 2018)

	Statement			Q&A Session		
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.122	0.115	0.141	0.031	0.072	-0.043
	(0.108)	(0.110)	(0.140)	(0.049)	(0.067)	(0.076)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.018**	0.004	0.006	0.006	0.001	-0.011
	(0.008)	(0.014)	(0.014)	(0.004)	(0.006)	(0.009)
TP_t^M		0.079	0.070		0.112	0.154
		(0.196)	(0.201)		(0.156)	(0.157)
TP_t^{FS}		-0.271	-0.242		-0.061	-0.188
		(0.241)	(0.259)		(0.184)	(0.193)
$Sentiment_t^M$		0.260**	0.243**		-0.018	0.057
		(0.115)	(0.115)		(0.078)	(0.088)
$Sentiment_t^{FS}$		-0.042	-0.037		0.074	0.053
		(0.053)	(0.057)		(0.046)	(0.042)
i_{t-1}			-0.007			0.030**
			(0.015)			(0.012)
$Constant$	-0.081	-0.099	-0.140	-0.071	-0.204	-0.022
	(0.185)	(0.260)	(0.301)	(0.082)	(0.156)	(0.141)
Sample	2011-2018	2011-2018	2011-2018	2011-2018	2011-2018	2011-2018
N	32	32	32	32	32	32
$Adj.R^2$	0.353	0.424	0.404	0.036	0.060	0.240

Note: The table reports OLS estimates of equation (9) for the period from April 2011 to December 2018. Each observation corresponds to one press conference. The dependent variable is the sentiment of the Chair's introductory statement in columns (1)-(3) and the sentiment of the Chair's responses during the Q&A session in columns (4)-(6). $\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$ ddenote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the unemployment gap, respectively. TP_t^M and TP_t^{FS} are the proportions of FOMC discussions recorded in the minutes devoted to mandate objectives and financial stability, while $Sentiment_t^M$ and $Sentiment_t^{FS}$ are the corresponding topic-specific sentiment measures. i_{t-1} denotes the policy rate at the immediately preceding FOMC meeting, measured by the effective federal funds rate and replaced by the shadow rate of Wu and Xia (2016) when the policy rate is at the ZLB. Sentiment is measured using the fine-tuned RoBERTa model described in Section 3.3. Bootstrapped standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table B11: Taylor Rule Estimates with Two-Quarter Forecast Averages of core PCE Inflation and the Output-Gap (2000–2018)

	FFR				Shadow Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Long-run estimates						
$\mathbb{E}_t\left(\bar{\pi}_t^{1:2}\right)$	0.928	1.420	2.542	4.712**	5.461	5.382
	(1.540)	(1.060)	(1.695)	(2.318)	(3.748)	(4.657)
$\mathbb{E}_t\left(\bar{y}_t^{1:2}\right)$	0.665*	0.606**	0.501	1.029	2.408	2.826
	(0.366)	(0.263)	(0.376)	(0.748)	(2.054)	(2.961)
Constant	1.154	8.065*	7.428	8.452	17.340	27.092
	(3.036)	(4.332)	(6.516)	(7.894)	(15.748)	(27.928)
Short-run estimates						
i_{t-1}	0.948***	0.929***	0.949***	0.949***	0.967***	0.974***
	(0.023)	(0.025)	(0.023)	(0.027)	(0.024)	(0.024)
$\mathbb{E}_t\left(\bar{\pi}_t^{1:2}\right)$	0.049	0.101	0.130	0.240**	0.180*	0.140
	(0.092)	(0.094)	(0.104)	(0.108)	(0.103)	(0.101)
$\mathbb{E}_t\left(\bar{y}_t^{1:2}\right)$	0.035***	0.043***	0.026	0.052**	0.080***	0.074***
	(0.011)	(0.015)	(0.018)	(0.025)	(0.025)	(0.026)
TP_t^M		-0.734**	-0.631*	-0.854*	-0.981**	-0.961**
		(0.297)	(0.325)	(0.451)	(0.425)	(0.428)
TP_t^{FS}		-0.698*	-0.369	-1.286**	-1.259**	-1.162**
		(0.361)	(0.337)	(0.548)	(0.511)	(0.500)
$Sentiment_t^M$			0.454***	0.725***	0.698***	0.640***
			(0.172)	(0.196)	(0.195)	(0.177)
$Sentiment_t^{FS}$			-0.032	-0.041	0.000	0.005
			(0.102)	(0.119)	(0.112)	(0.112)
VIX_t						-0.006
						(0.007)
D_t^{ZLB}				-0.518	-1.018***	-1.067***
				(0.329)	(0.392)	(0.395)
$D_t^{ZLB} \times TP_t^M$				0.686	1.398***	1.442***
				(0.443)	(0.498)	(0.498)
$D_t^{ZLB} \times TP_t^{FS}$				1.591**	2.522***	2.575***
				(0.701)	(0.797)	(0.790)
$D_t^{ZLB} \times Sentiment_t^M$				-0.926***	-0.817***	-0.806***
				(0.261)	(0.289)	(0.279)
$D_t^{ZLB} \times Sentiment_t^{FS}$				0.064	0.182	0.146
				(0.157)	(0.176)	(0.191)
Constant	0.060	0.574**	0.380	0.431	0.573	0.705*
	(0.144)	(0.266)	(0.321)	(0.402)	(0.390)	(0.407)
Wald Tests (χ^2)						
$\theta_{TP^M} + \varphi_{TP^M} = 0$				0.702	1.547	2.050
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$				0.873	5.926**	6.692**
$\theta_{Sent^M} + \varphi_{Sent^M} = 0$				1.090	0.307	0.569
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$				0.042	1.726	1.024
Sample	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018
N	152	152	152	152	152	152
$Adj.R^2$	0.970	0.972	0.973	0.976	0.984	0.984

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018. In this robustness check, the baseline averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap are replaced by averages over horizons one and two. Columns (1)–(4) use the effective federal funds rate as the dependent variable. Columns (5)–(6) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period, while column (6) additionally controls for the VIX. $\mathbb{E}_t\left(\bar{\pi}_t^{1:2}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:2}\right)$ denote the averages of the one- and two-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate and financial stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment measures, constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. Long-run estimates are obtained by dividing the corresponding short-run coefficients by $1-\rho$, where ρ denotes the estimated coefficient on the lagged dependent variable. The Wald test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table B12: Sentiment Regression with Two-Quarter Forecast Averages of core PCE Inflation and the Output Gap

	Statement			Q&A Session		
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{E}_t(\bar{\pi}_t^{1:2})$	0.145	0.153*	0.200*	0.020	0.046	-0.045
	(0.096)	(0.090)	(0.104)	(0.043)	(0.056)	(0.063)
$\mathbb{E}_t(\bar{y}_t^{1:2})$	0.017**	0.002	0.011	0.006	0.004	-0.012
	(0.007)	(0.012)	(0.013)	(0.004)	(0.006)	(0.009)
TP_t^M		0.078	0.039		0.084	0.159
		(0.186)	(0.190)		(0.149)	(0.152)
TP_t^{FS}		-0.299	-0.220		-0.044	-0.198
		(0.219)	(0.245)		(0.179)	(0.186)
$Sentiment_t^M$		0.266**	0.223**		-0.026	0.059
		(0.104)	(0.104)		(0.078)	(0.084)
$Sentiment_t^{FS}$		-0.035	-0.022		0.073	0.048
		(0.053)	(0.058)		(0.045)	(0.041)
i_{t-1}			-0.017			0.034***
			(0.015)			(0.013)
Constant	-0.110	-0.152	-0.211	-0.049	-0.140	-0.025
	(0.163)	(0.223)	(0.228)	(0.072)	(0.138)	(0.124)
Sample	2011-2018	2011-2018	2011-2018	2011-2018	2011-2018	2011-2018
N	32	32	32	32	32	32
$Adj.R^2$	0.383	0.483	0.484	0.044	0.057	0.250

Note: The table reports OLS estimates of equation (9) for the period from April 2011 to December 2018. Each observation corresponds to one press conference. The dependent variable is the sentiment of the Chair's introductory statement in columns (1)–(3) and the sentiment of the Chair's responses during the Q&A session in columns (4)–(6). $\mathbb{E}_t\left(\bar{\pi}_t^{1:2}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:2}\right)$ denote the averages of the one- and two-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. TP_t^M and TP_t^{FS} are the proportions of FOMC discussions devoted to mandate objectives and financial stability, while $Sentiment_t^M$ and $Sentiment_t^{FS}$ are the corresponding topic-specific sentiment measures. i_{t-1} denotes the policy rate at the previous FOMC meeting, measured by the effective federal funds rate and replaced by the shadow rate of Wu and Xia (2016) during the ZLB period. Sentiment is measured using the fine-tuned RoBERTa model described in Section 3.3. Bootstrapped standard errors are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is indicated by *, **, and *** respectively.

Table B13: Taylor Rule Estimations with Different Financial Indicators (2000-2018)

	FFR				Shadow Rate			
	VIX	NFCI	FSI	KCFSI	VIX	NFCI	FSI	KCFSI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Short-run estimates:								
i_{t-1}	0.958***	0.949***	0.959***	0.955***	0.976***	0.971***	0.977***	0.976***
	(0.028)	(0.027)	(0.027)	(0.028)	(0.025)	(0.025)	(0.024)	(0.026)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.219*	0.284**	0.221*	0.257**	0.164	0.203*	0.161	0.178
	(0.114)	(0.112)	(0.115)	(0.118)	(0.112)	(0.107)	(0.111)	(0.115)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.038	0.042*	0.037	0.042*	0.073***	0.077***	0.072***	0.075***
	(0.026)	(0.025)	(0.026)	(0.025)	(0.026)	(0.026)	(0.026)	(0.025)
TP_t^M	-0.855*	-0.870*	-0.877*	-0.867*	-0.991**	-1.010**	-1.003**	-0.993**
	(0.458)	(0.453)	(0.450)	(0.448)	(0.427)	(0.428)	(0.424)	(0.422)
TP_t^{FS}	-1.171**	-1.139**	-1.151**	-1.184**	-1.171**	-1.234**	-1.153**	-1.164**
	(0.544)	(0.528)	(0.527)	(0.534)	(0.511)	(0.513)	(0.499)	(0.507)
$Sentiment_t^M$	0.628***	0.659***	0.572***	0.638***	0.636***	0.676***	0.600***	0.640***
	(0.165)	(0.167)	(0.163)	(0.170)	(0.176)	(0.179)	(0.180)	(0.181)
$Sentiment_t^{FS}$	-0.033	-0.056	-0.029	-0.041	0.004	0.000	0.006	-0.001
	(0.115)	(0.109)	(0.114)	(0.114)	(0.110)	(0.106)	(0.110)	(0.109)
$Financial\ Indicator_t$	-0.009	-0.084	-0.039*	-0.044	-0.004	-0.009	-0.022	-0.024
	(0.006)	(0.085)	(0.022)	(0.039)	(0.007)	(0.104)	(0.026)	(0.045)
D_t^{ZLB}	-0.684**	-0.623*	-0.723**	-0.619*	-1.100***	-1.075***	-1.124***	-1.062***
	(0.336)	(0.335)	(0.332)	(0.335)	(0.397)	(0.399)	(0.398)	(0.400)
$D_t^{ZLB} \times TP_t^M$	0.809*	0.746*	0.825*	0.729*	1.485***	1.459***	1.497***	1.439***
	(0.445)	(0.446)	(0.440)	(0.443)	(0.497)	(0.504)	(0.497)	(0.501)
$D_t^{ZLB} \times TP_t^{FS}$	1.802***	1.732**	1.783***	1.766**	2.617***	2.597***	2.609***	2.589***
	(0.695)	(0.697)	(0.678)	(0.694)	(0.810)	(0.809)	(0.807)	(0.813)
$D_t^{ZLB} \times Sentiment_t^M$	-0.952***	-0.979***	-0.885***	-0.965***	-0.835***	-0.852***	-0.796***	-0.844***
	(0.247)	(0.259)	(0.230)	(0.249)	(0.278)	(0.292)	(0.269)	(0.280)
$D_t^{ZLB} \times Sentiment_t^{FS}$	0.016	0.059	0.027	0.040	0.152	0.178	0.156	0.163
	(0.167)	(0.162)	(0.160)	(0.164)	(0.192)	(0.188)	(0.189)	(0.190)
$Constant$	0.592	0.312	0.456	0.384	0.652*	0.531	0.588	0.546
	(0.386)	(0.399)	(0.385)	(0.386)	(0.396)	(0.393)	(0.380)	(0.379)
Wald Tests (χ^2)								
$\theta_{TP^M} + \varphi_{TP^M} = 0$	0.053	0.401	0.075	0.503	2.167	1.793	2.208	1.738
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$	2.999*	2.369	3.420*	2.553	6.810***	5.584**	6.993***	6.440**
$\theta_{Sent^M} + \varphi_{Sent^M} = 0$	2.647	2.328	2.773*	2.461	0.811	0.575	0.832	0.801
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$	0.021	0.000	0.000	0.000	1.056	1.436	1.131	1.144
Sample	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018
N	152	152	152	152	152	152	152	152
$Adj.R^2$	0.977	0.977	0.978	0.977	0.984	0.984	0.984	0.984

Note: The table shows the Taylor rule estimation results, derived by the OLS regression of equation (6), covering the period from January 2000 to December 2018. Each specification includes one of four financial indicators: the CBOE Volatility Index (VIX), the Chicago Fed National Financial Conditions Index (NFCI), the St. Louis Fed Financial Stress Index (FSI), or the Kansas City Financial Stress Index (KCFSI). Each indicator is measured as its average value between consecutive FOMC meetings. Columns (1)–(4) use the effective federal funds rate as the dependent variable, whereas columns (5)–(8) use the shadow rate measure of Wu and Xia (2016) during the ZLB period. $\mathbb{E}_t(\bar{\pi}_t^{1:4})$ and $\mathbb{E}_t(\bar{y}_t^{1:4})$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively, while i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate and financial stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment measures, classified using the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. The Wald-test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table B14: Taylor Rule Estimations with Second-Order Interest Rate Smoothing (2000-2018)

	FFR				Shadow Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Long-run estimates						
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	1.322	1.887*	3.092*	5.820**	6.912	6.846
	(1.367)	(1.025)	(1.751)	(2.560)	(5.320)	(7.999)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.670*	0.590**	0.478	0.921	2.969	4.264
	(0.361)	(0.256)	(0.370)	(0.706)	(2.886)	(5.990)
Constant	0.440	7.285**	6.519	6.214	18.585	38.161
	(2.583)	(3.619)	(5.534)	(6.124)	(17.426)	(50.176)
Short-run estimates						
i_{t-1}	0.882***	0.857***	0.864***	0.808***	0.784***	0.770***
	(0.161)	(0.158)	(0.162)	(0.155)	(0.132)	(0.142)
i_{t-2}	0.064	0.070	0.083	0.137	0.186	0.210
	(0.163)	(0.155)	(0.157)	(0.152)	(0.131)	(0.140)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.071	0.136	0.161	0.320***	0.206*	0.134
	(0.085)	(0.096)	(0.109)	(0.117)	(0.110)	(0.113)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.036***	0.043***	0.025	0.051**	0.088***	0.084***
	(0.013)	(0.015)	(0.018)	(0.025)	(0.027)	(0.027)
TP_t^M		-0.755***	-0.649**	-0.891**	-1.025**	-0.987**
		(0.292)	(0.311)	(0.418)	(0.400)	(0.396)
TP_t^{FS}		-0.698*	-0.375	-1.417**	-1.335***	-1.198**
		(0.381)	(0.354)	(0.562)	(0.516)	(0.493)
$Sentiment_t^M$			0.450***	0.748***	0.714***	0.643***
			(0.169)	(0.188)	(0.190)	(0.175)
$Sentiment_t^{FS}$			-0.024	-0.035	-0.002	-0.001
			(0.101)	(0.117)	(0.111)	(0.109)
VIX_t						-0.007
						(0.006)
D_t^{ZLB}				-0.593*	-1.076***	-1.113***
				(0.322)	(0.404)	(0.420)
$D_t^{ZLB} \times TP_t^M$				0.742*	1.499***	1.538***
				(0.423)	(0.500)	(0.516)
$D_t^{ZLB} \times TP_t^{FS}$				1.832**	2.705***	2.737***
				(0.720)	(0.849)	(0.844)
$D_t^{ZLB} \times Sentiment_t^M$				-1.019***	-0.961***	-0.950***
				(0.273)	(0.311)	(0.302)
$D_t^{ZLB} \times Sentiment_t^{FS}$				0.088	0.283	0.252
				(0.169)	(0.196)	(0.203)
Constant	0.024	0.526**	0.340	0.341	0.554	0.747*
	(0.134)	(0.239)	(0.284)	(0.345)	(0.352)	(0.402)
Wald Tests (χ^2)						
$\rho_1 + \rho_2 = 1$	5.900**	8.295***	5.100**	3.663*	1.381	0.608
$\theta_{TP^M} + \varphi_{TP^M} = 0$				0.535	1.991	2.556
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$				1.554	6.100**	6.838***
$\theta_{Sent^M} + \varphi_{Sent^M} = 0$				1.807	1.078	1.587
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$				0.187	3.304*	2.441
Sample	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018
N	152	152	152	152	152	152
$Adj.R^2$	0.971	0.972	0.973	0.977	0.984	0.985

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018. As a robustness check, the baseline specification is extended to allow for second-order interest rate smoothing by including two lags of the dependent variable. Columns (1)–(4) use the effective federal funds rate as the dependent variable. Columns (5)–(6) replace it with the shadow rate measure of Wu and Xia (2016) during the ZLB period, while column (6) additionally controls for the VIX. $\mathbb{E}_t(\bar{\pi}_t^{1:4})$ and $\mathbb{E}_t(\bar{y}_t^{1:4})$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} and i_{t-2} denote the first and second lag of the dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate and financial stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment measures, constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. Long-run estimates are obtained by dividing the corresponding short-run coefficients by $1 - \rho_1 - \rho_2$, where ρ_1 and ρ_2 denote the estimated coefficients on the first and second lag of the dependent variable. The Wald test rows report χ^2 statistics for the unit-root restriction on the interest-rate smoothing parameters and for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table B15: Taylor Rule Estimates Using the Krippner Shadow Rate (2000–2018)

	FFR				Shadow Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
Long-run estimates						
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	1.320	1.875*	3.039*	5.619**	4.432	3.838
	(1.378)	(1.027)	(1.779)	(2.410)	(2.745)	(3.085)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.653*	0.582**	0.465	0.849	1.874*	2.027
	(0.343)	(0.254)	(0.374)	(0.646)	(1.047)	(1.318)
Constant	0.441	7.401*	6.748	6.686	14.542	21.710
	(2.635)	(3.935)	(6.098)	(6.860)	(9.319)	(13.623)
Short-run estimates						
i_{t-1}	0.947***	0.929***	0.949***	0.945***	0.955***	0.963***
	(0.022)	(0.025)	(0.022)	(0.027)	(0.020)	(0.021)
$\mathbb{E}_t(\bar{\pi}_t^{1:4})$	0.069	0.133	0.156	0.309***	0.198*	0.143
	(0.085)	(0.095)	(0.109)	(0.118)	(0.115)	(0.115)
$\mathbb{E}_t(\bar{y}_t^{1:4})$	0.034***	0.041***	0.024	0.047*	0.084***	0.076**
	(0.012)	(0.014)	(0.018)	(0.025)	(0.027)	(0.029)
TP_t^M		-0.755**	-0.651**	-0.914**	-1.093**	-1.051**
		(0.300)	(0.325)	(0.448)	(0.438)	(0.446)
TP_t^{FS}		-0.682*	-0.367	-1.358**	-1.358**	-1.234**
		(0.368)	(0.344)	(0.547)	(0.527)	(0.530)
$Sentiment_t^M$			0.438**	0.713***	0.625***	0.558***
			(0.171)	(0.189)	(0.193)	(0.170)
$Sentiment_t^{FS}$			-0.026	-0.033	-0.020	-0.020
			(0.102)	(0.117)	(0.113)	(0.111)
VIX_t						-0.007
						(0.006)
D_t^{ZLB}				-0.630*	-0.559	-0.580
				(0.334)	(0.532)	(0.535)
$D_t^{ZLB} \times TP_t^M$				0.786*	0.431	0.428
				(0.444)	(0.701)	(0.697)
$D_t^{ZLB} \times TP_t^{FS}$				1.783**	3.220***	3.234***
				(0.719)	(0.925)	(0.923)
$D_t^{ZLB} \times Sentiment_t^M$				-0.963***	0.484	0.495
				(0.261)	(0.410)	(0.403)
$D_t^{ZLB} \times Sentiment_t^{FS}$				0.060	-0.128	-0.161
				(0.158)	(0.221)	(0.233)
Constant	0.023	0.526**	0.347	0.368	0.650	0.810**
	(0.134)	(0.249)	(0.308)	(0.383)	(0.402)	(0.407)
Wald Tests (χ^2)						
$\theta_{TP^M} + \varphi_{TP^M} = 0$				0.429	1.260	1.104
$\theta_{TP^{FS}} + \varphi_{TP^{FS}} = 0$				1.597	7.280***	7.637***
$\theta_{Sent^M} + \varphi_{Sent^M} = 0$				1.586	9.912***	8.574***
$\theta_{Sent^{FS}} + \varphi_{Sent^{FS}} = 0$				0.059	0.573	0.762
Sample	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018	2000-2018
N	152	152	152	152	152	152
$Adj.R^2$	0.971	0.972	0.973	0.977	0.983	0.984

Note: The table reports OLS estimates of equation (6) for the period from January 2000 to December 2018. The specification replaces the Wu–Xia shadow rate used in the baseline analysis with the Krippner shadow rate estimates (Krippner, 2013). Columns (1)–(4) use the effective federal funds rate as the dependent variable. Columns (5)–(6) replace it with the Krippner shadow rate during the ZLB period, while column (6) additionally controls for the VIX. $\mathbb{E}_t(\bar{\pi}_t^{1:4})$ and $\mathbb{E}_t(\bar{y}_t^{1:4})$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate and financial-stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment measures, constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. D_t^{ZLB} equals one from December 2008 to December 2015 and zero otherwise. Long-run estimates are obtained by dividing the corresponding short-run coefficients by $1-\rho$, where ρ denotes the estimated coefficient on the lagged dependent variable. The Wald test rows report χ^2 statistics for the total effects during the ZLB period. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

B.6. Taylor Rule Estimates by Monetary Policy Regime

Table B16: Subsample Analysis: Taylor Rule Estimates by Monetary Policy Regime

	Non-ZLB	ZLB	ZLB excluding GFC
	(1)	(2)	(3)
Short-run estimates			
i_{t-1}	0.891***	1.091***	1.113***
	(0.030)	(0.074)	(0.079)
$\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$	0.539***	0.078	0.023
	(0.148)	(0.140)	(0.124)
$\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$	0.155***	0.065***	0.066***
	(0.049)	(0.021)	(0.024)
TP_t^M	-1.884***	0.588*	0.685**
	(0.559)	(0.318)	(0.334)
TP_t^{FS}	-2.231***	1.022***	0.819**
	(0.593)	(0.386)	(0.404)
$Sentiment_t^M$	0.369*	0.122	0.208
	(0.201)	(0.159)	(0.221)
$Sentiment_t^{FS}$	0.095	0.110	0.098
	(0.097)	(0.110)	(0.113)
$Constant$	0.825*	-0.260	-0.168
	(0.463)	(0.344)	(0.376)
Sample period	2000.01–2008.11 2016.01–2018.12	2008.12–2015.12	2010.01–2015.12
N	95	57	48
Adj. R^2	0.970	0.944	0.934

Note: The table reports OLS estimates of a subsample version of equation (6), estimated separately across monetary policy regimes. Because the specification is estimated separately by regime, the ZLB indicator and its interactions with the explanatory variables are omitted. Column (1) covers the non-ZLB periods from January 2000 to November 2008 and from January 2016 to December 2018. Column (2) covers the full ZLB period from December 2008 to December 2015. Column (3) restricts the ZLB sample to January 2010 through December 2015, thereby excluding the GFC observations at the beginning of the ZLB period. The dependent variable for Column (1) is the effective federal funds rate and for Column (2) and Column (3) the Wu and Xia (2016) shadow rate estimates. $\mathbb{E}_t\left(\bar{\pi}_t^{1:4}\right)$ and $\mathbb{E}_t\left(\bar{y}_t^{1:4}\right)$ denote the averages of the one- through four-quarter-ahead Greenbook forecasts of core PCE inflation and the output gap, respectively. i_{t-1} denotes the lagged dependent variable. TP_t^M and TP_t^{FS} denote the topic proportions of mandate-related and financial stability discussions, respectively. $Sentiment_t^M$ and $Sentiment_t^{FS}$ denote the corresponding sentiment measures, constructed from sentence-level classifications generated by the fine-tuned RoBERTa model described in Section 3.3. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

C. Quotes about Financial Stability Considerations in the FOMC Transcripts

To complement the econometric evidence, we conduct a manual narrative analysis of the FOMC transcripts inspired by Romer and Romer (2023). We identify passages in which participants explicitly relate financial stability considerations to the appropriate stance or timing of monetary policy. The excerpts illustrate two distinct channels. During periods of acute financial stress, financial stability concerns supported monetary accommodation, whereas concerns about excessive risk-taking, reaching for yield, and asset price imbalances provided arguments for an earlier withdrawal of accommodation. The narrative evidence is intended to contextualize the econometric results rather than to provide an independent identification strategy.

Table C1: Illustrative FOMC Transcript Excerpts on Financial Stability and Monetary Policy

Excerpts	**Meeting Date**	**FOMC Member**	**Quote**
1	September 18, 2007	Eric S. Rosengren	"So we have a situation of a very weak housing sector, some evidence of slowing employment growth, and a period of extended illiquidity that may get worse before it gets better. The tail risk of liquidity problems and economic problems has grown, and we clearly want to avoid outcomes by which declines in prices for houses and for financial assets tied to the housing sector could create more-severe economic outcomes. The fact is that we do not have much experience with periods of extended illiquidity, especially when the housing sector is so weak. So taking out insurance against these risks seems entirely appropriate. The decision is made easier, in my view, because I see the risk of unacceptably high inflation resulting from such an action as being quite low. My hope is that with appropriate easing of policy the liquidity issues will abate as we start the fourth quarter."
2	March 16, 2010	Thomas M. Hoenig	"The reason to begin tightening policy soon is not simply or even primarily to prevent the issues around inflationary pressures immediately. That's longer term and important. Rather it is to prevent future buildup of financial imbalances and reduce the risks to longer-run macroeconomic and financial stability that come from maintaining a commitment to exceptionally low rates for an extended period of time."
3	October 23-24, 2012	Daniel K. Tarullo	"I have concerns about more purchases. As others have pointed out, the dealer community is now assuming close to a $4 trillion balance sheet and purchases through the first quarter of 2014. I admit that is a much stronger reaction than I anticipated, and I am uncomfortable with it for a couple of reasons. [...] Second, I think we are actually at a point of encouraging risk-taking, and that should give us pause. Investors really do understand now that we will be there to prevent serious losses. It is not that it is easy for them to make money but that they have every incentive to take more risk, and they are doing so. Meanwhile, we look like we are blowing a fixed-income duration bubble right across the credit spectrum that will result in big losses when rates come up down the road. You can almost say that that is our strategy."
4	April 30 - May 01, 2013	Esther L. George	"While I do support accommodative policy, continuing this policy well beyond our response to the crisis and recession, I think, poses risk to long-term sustainable economic growth. And the risk associated with zero interest rates for many years, combined with the ongoing expansion of our balance sheet, has become more significant. Evidence of financial imbalances has emerged, although, as many have noted, near-term risks to broad financial stability seem unlikely right now. [...] As risks surface in multiple areas, they individually may not threaten the broad economy, but they are symptomatic of conditions that give rise to excessive risk-taking, reaching for yield, and asset mispricing. Zero interest rates for an extended period of time create such conditions and, as such, pose real risk of painful future adjustments"

(continued)

Excerpts	Meeting Date	FOMC Member	Quote
5	January 28-29, 2014	Narayana Kocherlakota	"In the meantime, our policy settings continue to facilitate incentives that pose potential risks to long-term financial stability. Asset prices and froth in various pockets of the financial sector deserve our ongoing monitoring, particularly in high-yield bond issuance and leveraged lending. I recognize that identifying these risks to financial stability and quantifying their impact today remain difficult for sure, but we should not underestimate their potential effects or their ability to surprise."
6	March 18-19, 2014	Esther L. George	"That said, I would highlight a few issues that continue to concern me about this policy stance. [...] Second, I remain concerned about the side effects of unusually low interest rates for this extended period of time and the incentives they promote for risk-taking and reaching for yield. I continue to hear from contacts about fierce competition among lenders resulting in questionable terms and underwriting standards."
7	April 29-30, 2014	Janet L. Yellen	"Turning to the topic of what could throw our forecast off course, many of you mentioned the main risks that concern you as being external [...]; and financial-stability risks in that we're seeing evidence of reach-for-yield behavior in fixed-income markets, including leveraged loan markets, in which terms and standards are continuing to deteriorate in spite of our guidance. I thought Governor Stein's suggestion of looking at catastrophe bonds and the fact that their yields have moved with these spreads might lead us to worry a little bit more than we might have if we thought declines in spreads were reflecting greater health in the corporate sector. This suggests that it is a measure of risk appetite, and we should worry about it because there could come an unwinding at some point or a volatility event that would involve financial-stability risks. In particular, as the time of liftoff draws closer, we could see financial conditions tighten sharply."
8	June 17-18, 2014	Jerome H. Powell	"After almost six years of highly-accommodative policy, the risks are out there and continue to build. To be specific, the risk in the near term, to me, is not that of a meltdown at the core of the system, which is much stronger and better protected than before the crisis, but rather that of a sharp correction amplified by the liquidity mismatch in the markets that would damage or halt the progress of what is still a weak economy. As we've discussed, and I would be in agreement generally with Bill's approach to this question, but as we discussed yesterday morning with the staff, implied volatility for equities, currencies, and long- and short-term fixed income are at historically low levels around the globe and market participants have pointed to a number of explanations for that. Particularly important for me would be, along with low realized volatility and the reach for yield, the guidance that this Committee and other central banks have given, that rates will remain low for the foreseeable future. This explanation, to me, cuts across and supports all of the other explanations."
9	October 28-29, 2014	Stanley Fischer	"On financial stability, that's a very tough one indeed. The current constitutional document is based in the legal system, in laws relating to our mandate. I don't think financial stability is equally based. There's a slight problem there. Nonetheless, I liked a lot of what I heard around the table. I think that the key issue here is not how the Fed deals with financial stability but how the FOMC deals with financial stability. And that means we control the interest rate, and do we ever use the interest rate to deal with financial-stability issues? Now, President Evans used the Greenspan line: "Well, if you were just going to use the interest rate when the situation becomes untenable, you'd have to do so much to raise the interest rate that you'd destroy the economy." That prejudged the answer. We can't do it. We wait for the bubble to pop, and then we mop it up. The issue is, do we take financial-stability considerations into account in our interest rate decisions? We sometimes do, and then Lael pointed out that sometimes we do it right at the beginning and not right at the end. I think we ought to work a little bit more on that, but we cannot develop the statement that implies it's on the same level as inflation and unemployment."

(continued)

Excerpts	Meeting Date	FOMC Member	Quote
10	December 16-17, 2014	Daniel K. Tarullo	"The final thing I wanted to say is, the analytic issue of whether the greater risk lies in falling behind an inflation increase, on the one hand, or a premature liftoff, on the other, is one to which many people have alluded, but we haven't actually discussed it very much. People have more given their conclusions, either based on faith or based on some reference to past circumstances. I'm inclined to think that the risk of premature liftoff in current circumstances is greater than the risk of falling behind the inflation curve. In fact, as I sit here today, I think it's at least as likely that I'll ultimately be motivated to favor liftoff for financial-stability reasons as for inflation reasons."
11	April 28-29, 2015	Narayana Kocherlakota	"As I've indicated in the past, I believe that it is important for monetary policy decisions to take account of the macroeconomic risks created by potential financial instability. My own reading of the QS report is that the main financial stability risk we face right now is actually associated with higher, not lower, interest rates. This will build on some of the work that Thomas was showing us earlier this morning."
12	October 27-28, 2015	Janet L. Yellen	"I want to start by reviewing how, in my view, the economic situation has evolved over the past few months. At the time of our July meeting, my expectation was that if the federal funds rate remained near zero, monthly payroll gains would remain close to 200,000, and the unemployment rate would steadily decline. Delaying liftoff until the transitory effects of low oil prices and dollar appreciation on 12-month inflation had totally dissipated—something that looked quite unlikely before spring of 2016-risked a situation in which the unemployment rate might well hit $4\frac{1}{2}$ percent by the middle of next year, with further declines quite likely. I worried that excessive pressure on resources could well develop under these conditions, eventually generating an unwelcome rise in inflation or inappropriate risk-taking in financial markets. To avoid ending up in a situation in which we might need to tighten policy abruptly— something that could easily disrupt the economy-I anticipated that it would be appropriate to start tightening at our September meeting."

Note: The table reports selected excerpts from official FOMC transcripts in which participants relate financial stability considerations to the stance or timing of monetary policy. The excerpts are illustrative rather than exhaustive and complement the econometric analysis. Bracketed ellipses indicate omitted text.

D. Model Fine-tuning Approach, Evaluation, and Comparison

To fine-tune the LLMs, we created a dataset consisting of 2000 randomly selected sentences from the discussion sections of the Minutes spanning the years 1993 to 2022. The first two authors independently classified the sentences into three sentiment categories, namely Negative, Neutral and Positive, from the perspective of the Federal Reserve. In cases of disagreement the third author reviewed the classifications and made the final decision. To ensure uniform classification, each sentence was evaluated based solely on the information contained within it. This approach implies that sentences about rising or falling inflation rates are classified as positive or negative only if it is explicitly clear that the direction of the inflation rate aligns with the mandate. Otherwise, these sentences are classified as neutral. After classification, the dataset comprised 793 neutral sentences, 648 negative sentences, and 559 positive sentences. This dataset was used to fine-tune two different LLMs: FinBERT and RoBERTa. The labeled dataset is available in our Google Drive Online Repository (URL).

The fine-tuning process for the RoBERTa model is described here, with the same approach applied to the FinBERT model. The dataset was split into training (70%), validation (15%), and test (15%) sets using stratified sampling to preserve the distribution of sentiment labels across the splits. A pre-trained RoBERTa model (`roberta-base`) was selected for adaptation to classify sentences into three sentiment classes. The tokenizer and model were obtained from the Hugging Face library. The text data was tokenized using the `RobertaTokenizer` with a maximum sequence length of 512 tokens. Padding and truncation were applied to ensure uniform sequence lengths compatible with the RoBERTa model's input requirements.

A comprehensive grid search was conducted across multiple combinations of hyperparameters, with a particular focus on optimizing the learning rate and batch size. A range of values for the learning rate was tested, from 10^{-7} to 10^{-4}. In addition, five values for the batch size were tested: 16, 32, 64, 128 and 256. Each hyperparameter combination was subjected to a systematic evaluation using a weighted F1 score and accuracy on the validation dataset, thereby ensuring a comprehensive search for the optimal configuration. An evaluation was conducted at the conclusion of each epoch to ascertain the accuracy and F1 score on the validation set. The final evaluation metrics were logged, and the best model was selected based on the lowest validation loss and highest F1 score.

Once the optimal hyperparameters were identified, the model was retrained using the best settings: a learning rate of 3×10^{-5}, a batch size of 16, and 3 epochs, with 50 warm-up steps. To maximize training data, the training and validation sets were merged, and the model was fine-tuned again. The final fine-tuned model was saved for subsequent evaluation on the test set. The performance of the model was then assessed on the test dataset. The results of the model performance, based on the 300 test sentences, are presented in Table D1.

Table D1: Classification Performance and Selected Hyperparameters

	Fine-tuned RoBERTa			
	Precision	Recall	F1-Score	Support
Neutral	0.83	0.71	0.77	119
Positive	0.73	0.79	0.75	84
Negative	0.80	0.88	0.84	97
Accuracy			0.79	300
Macro Avg.	0.78	0.79	0.79	300
Weighted Avg.	0.79	0.79	0.79	300

	Fine-tuned FinBERT			
	Precision	Recall	F1-Score	Support
Neutral	0.81	0.83	0.82	119
Positive	0.69	0.65	0.67	84
Negative	0.70	0.71	0.71	97
Accuracy			0.74	300
Macro Avg.	0.73	0.73	0.73	300
Weighted Avg.	0.74	0.74	0.74	300

	FinBERT Base Model			
	Precision	Recall	F1-Score	Support
Neutral	0.79	0.71	0.75	119
Positive	0.70	0.57	0.63	84
Negative	0.62	0.80	0.70	97
Accuracy			0.70	300
Macro Avg.	0.70	0.69	0.69	300
Weighted Avg.	0.71	0.70	0.70	300

Model	Learning Rate	Batch Size	Epochs
FinBERT	4×10^{-5}	256	9
RoBERTa	3×10^{-5}	16	3

Note: The table reports the classification performance of the fine-tuned RoBERTa model, the fine-tuned FinBERT model, and the FinBERT base model. Precision measures the proportion of predictions assigned to a given class that are correct, whereas recall measures the proportion of observations in that class that are correctly identified. The F1-score is the harmonic mean of precision and recall. Support denotes the number of manually labeled sentences in each sentiment category. The test sample comprises 300 sentences. Accuracy measures the overall proportion of correct predictions. The macro average assigns equal weight to each sentiment category, whereas the weighted average accounts for differences in class size. The final panel reports the learning rate, batch size, and number of epochs used to fine-tune the RoBERTa and FinBERT models.

In the next step, we compare our fine-tuned RoBERTa model with lexicon-based methods, which remain among the most widely used approaches in this literature. Although the dictionary developed by Loughran and McDonald (2011) was originally designed for the analysis of corporate disclosures, it is also frequently applied to central bank communication. Table D2 compares the classification performance of the LM dictionary and our fine-tuned RoBERTa model using the same manually labeled evaluation sample. The 300 sentences in this held-out sample were not used to fine-tune the RoBERTa model. The fine-tuned RoBERTa model achieves higher precision, recall, and F1-scores across all three sentiment categories and also substantially outperforms the dictionary in terms of overall accuracy.

Table D2: Classification Performance on the General Held-Out Evaluation Sample

	Loughran–McDonald Dictionary					Correa et al. (2021) Financial Stability Dictionary			
	Precision	Recall	F1-Score	Support		Precision	Recall	F1-Score	Support
Neutral	0.55	0.55	0.55	119	Neutral	0.49	0.81	0.61	119
Positive	0.52	0.43	0.47	84	Positive	0.58	0.33	0.42	84
Negative	0.54	0.63	0.58	97	Negative	0.64	0.37	0.47	97
Accuracy			0.54	300	Accuracy			0.53	300
Macro Avg.	0.54	0.53	0.53	300	Macro Avg.	0.57	0.50	0.50	300
Weighted Avg.	0.54	0.54	0.54	300	Weighted Avg.	0.57	0.53	0.51	300

	Fine-Tuned RoBERTa			
	Precision	Recall	F1-Score	Support
Neutral	0.83	0.71	0.77	119
Positive	0.73	0.79	0.75	84
Negative	0.80	0.88	0.84	97
Accuracy			0.79	300
Macro Avg.	0.78	0.79	0.79	300
Weighted Avg.	0.79	0.79	0.79	300

Note: The table compares the classification performance of the dictionary by Loughran and McDonald (2011), the financial stability dictionary developed by Correa et al. (2021), and the fine-tuned RoBERTa model on the same general held-out evaluation sample. The sample comprises 300 manually labeled sentences that were not used to fine-tune the RoBERTa model. Precision denotes the proportion of predictions assigned to a given sentiment category that are correct, whereas recall denotes the proportion of observations in that category that are correctly identified. The F1-score is the harmonic mean of precision and recall. Support reports the number of manually labeled sentences in each category. Accuracy measures the overall proportion of correctly classified sentences. The macro average assigns equal weight to each sentiment category, whereas the weighted average accounts for differences in class size. Higher values indicate better classification performance.

Finally, we compare our approach with the financial stability dictionary developed by Correa et al. (2021). This dictionary provides a particularly relevant benchmark because it was explicitly designed to measure sentiment in financial stability-related texts. A comparison based on our general evaluation sample also favors the fine-tuned RoBERTa model. However, this comparison may not provide the most appropriate assessment of the Correa et al. (2021) dictionary because the general evaluation sample contains many sentences that are not directly related to financial stability.

To provide a more targeted benchmark, we therefore draw an additional sample of 300 sentences from paragraphs classified as financial stability-related by the STM, considering only those sentences that we manually identify as explicitly addressing financial stability. None of these sentences is included in the data used to fine-tune the RoBERTa model. We manually classify the sentences as positive, neutral, or negative following the same annotation procedure used for the original evaluation dataset. Table D3 compares the performance of the two approaches on this financial stability-specific evaluation sample.

Table D3: Classification Performance on Financial Stability-Related Sentences

	Correa et al. (2021) Financial Stability Dictionary					Fine-Tuned RoBERTa Model			
	Precision	Recall	F1-Score	Support		Precision	Recall	F1-Score	Support
Neutral	0.60	0.65	0.62	116	Neutral	0.91	0.55	0.69	116
Positive	0.55	0.37	0.44	70	Positive	0.68	0.89	0.77	70
Negative	0.66	0.74	0.69	114	Negative	0.77	0.94	0.85	114
Accuracy			0.62	300	Accuracy			0.78	300
Macro Avg.	0.60	0.58	0.59	300	Macro Avg.	0.79	0.79	0.77	300
Weighted Avg.	0.61	0.62	0.61	300	Weighted Avg.	0.81	0.78	0.77	300

Note: The table compares the classification performance of the financial stability dictionary developed by Correa et al. (2021) and the fine-tuned RoBERTa model on a manually labeled sample of 300 financial-stability-related sentences. These sentences were drawn from the financial stability discussions identified in the FOMC minutes and were not included in the data used to fine-tune the RoBERTa model. Precision denotes the proportion of predictions assigned to a given category that are correct, while recall denotes the proportion of observations in that category that are correctly identified. The F1-score is the harmonic mean of precision and recall. Support reports the number of manually labeled sentences in each sentiment category. Accuracy measures the overall proportion of correctly classified sentences. The macro average assigns equal weight to each sentiment category, whereas the weighted average accounts for differences in class size. Higher values indicate better classification performance.

The fine-tuned RoBERTa model also outperforms the dictionary by Correa et al. (2021) in this more demanding comparison. It achieves higher F1-scores for all three sentiment categories and increases overall classification accuracy from 0.62 to 0.78. These results indicate that the fine-tuned RoBERTa model provides a consistent and more accurate framework for measuring sentiment across both mandate-related and financial stability-related discussions, without requiring different dictionaries for different economic topics. We provide the labeled dataset with annotated financial stability sentences also in our Google Drive Online Repository (URL).

E. Explainable AI and Transparent Sentiment Indicator

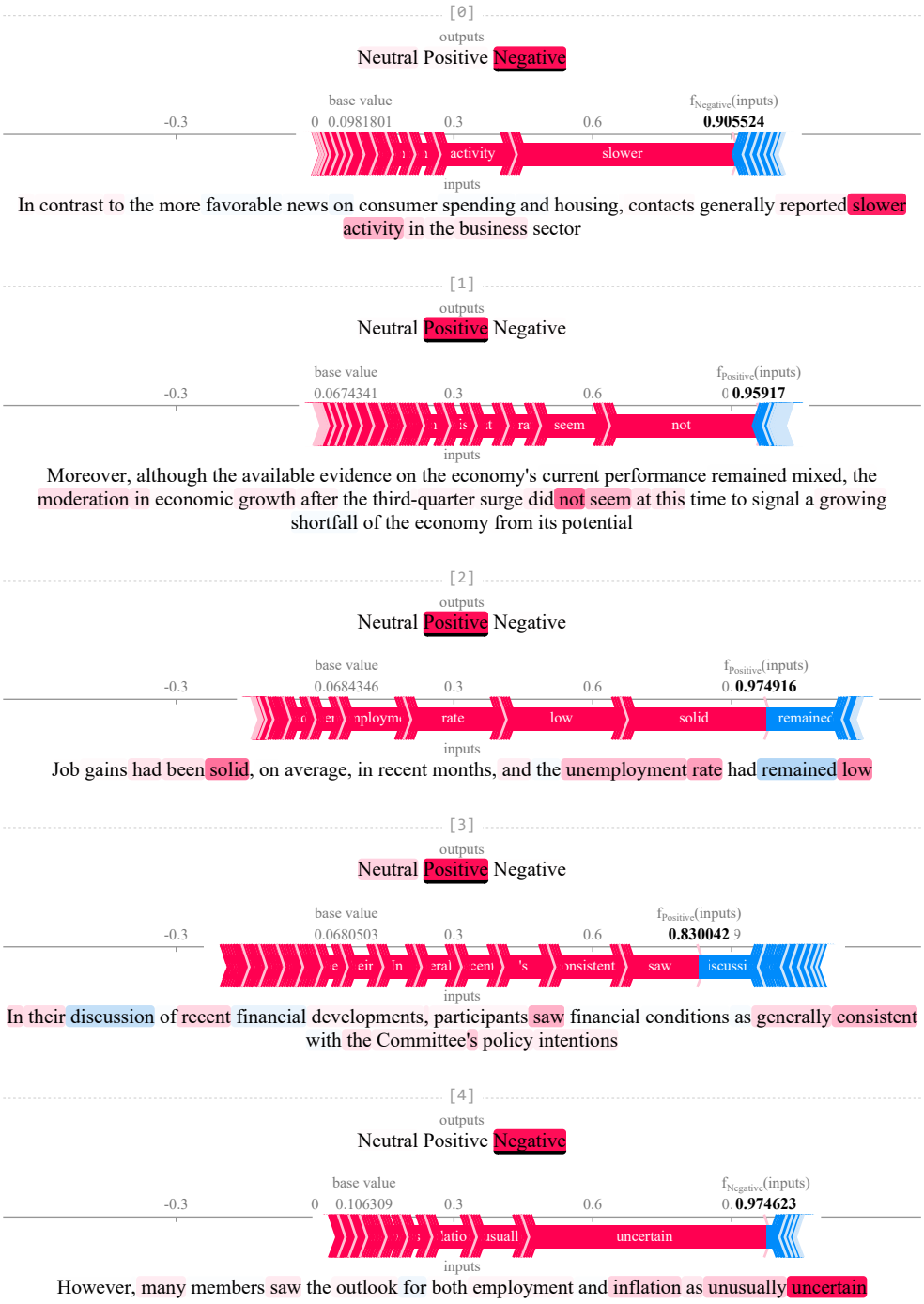
To examine how the fine-tuned RoBERTa model arrives at its sentiment classifications, we use SHAP to identify the contribution of individual words to each prediction (Lundberg and Lee, 2017). RoBERTa internally divides words into smaller subword units, but we define complete words as the explanatory features because they provide a more intuitive interpretation. The classifier itself continues to operate on its original subword inputs. SHAP only changes the level at which the prediction is explained.

For each sentence, SHAP creates alternative versions in which different words are masked and recalculates the model's predicted probability for a given sentiment class. A word receives a positive contribution when retaining it raises that probability and a negative contribution when retaining it lowers the probability. These comparisons are combined into a SHAP value for each word. The base value represents the model's predicted probability when all words are masked. Adding the word-level contributions to this base value yields the predicted probability for the complete sentence.

Figure E1 illustrates the decomposition for five sentences from the FOMC minutes. The model's three possible classifications (neutral, positive, negative) are shown above each sentence. The class whose prediction is being explained is underlined, while the shading of the class labels indicates the corresponding predicted probabilities. Red words increase the probability of the selected class, whereas blue words reduce it. Darker shading and wider arrow segments indicate larger contributions. The arrows show how the individual word contributions move the prediction from the base value to the probability assigned to the complete sentence.

In Section E.1, we provide the dictionary that replicates our RoBERTa model's sentiment measurement and present additional details on its derivation. We also evaluate it and compare it with the dictionary by Loughran and McDonald (2011).

Figure E1: Examples of SHAP Value Impact



Note: The figure illustrates the influence of individual words on sentiment classification for various sentences, as determined by our fine-tuned RoBERTa model. For each sentence, we display the calculation only for the sentiment category to which the sentence was assigned. A sentence is classified into the category with the highest probability score. The base value represents the probability of a sentence being classified into a sentiment category when all words are masked. The contributions of each word are then added to this base value, yielding the $f_{LABEL_i}(inputs)$ value. The intensity of the color reflects the contribution of each word, with more intense colors indicating a greater contribution. Words highlighted in red increase the probability of the assigned class, while those in blue decrease it.

E.1. A SHAP-Based Dictionary for Explainable Sentiment Analysis

We construct the positive and negative dictionaries from word-level SHAP explanations of the fine-tuned RoBERTa sentiment classifier. For each word occurrence, SHAP measures its contribution to the model's positive, negative, and neutral predictions, conditional on the sentence in which it appears. We first remove punctuation, non-alphabetic tokens, stop words, words with fewer than three characters, and irregular hyphenated expressions. The remaining words are then lemmatized, grouping different inflected forms under a common dictionary entry. For each lemma, we calculate the average SHAP contribution to the positive, negative, and neutral sentiment classes across all of its occurrences.

To reduce the influence of infrequent expressions, we retain only lemmas that appear at least ten times in the corpus. A lemma is classified as positive if its average contribution to the positive class exceeds its contributions to both the negative and neutral classes. We classify negative lemmas analogously. Within each group, we retain only the top 20% of lemmas ranked by their average SHAP contribution. The resulting word lists are generated automatically without manual adjustments. After preprocessing and applying the minimum-frequency requirement, inclusion depends solely on the SHAP values and the predefined percentile threshold. This procedure produces a dictionary containing 102 positive lemmas and 107 negative lemmas.

Because the dictionary is applied to words in their original form rather than to lemmatized text, we expand each selected lemma into its most common grammatical inflections. For verbs, these include the base form, third-person singular present, past tense, present participle, and past participle. For nouns, we include the singular and plural forms, while for adjectives we include the base, comparative, and superlative forms. Each inflected form inherits the sentiment classification of its underlying lemma. The resulting word list forms the SHAP dictionary, which is applied using exact word matching. This allows the dictionary to be used in the same way as established dictionary-based approaches, such as the dictionaries of Loughran and McDonald (2011) and Correa et al. (2021).

Tables E1 and E2 report the positive and negative word forms included in the SHAP-based sentiment dictionary, respectively. Table E3 compares the SHAP dictionary with the dictionary by Loughran and McDonald (2011) on 300 manually labeled sentences. The SHAP dictionary performs slightly better overall, achieving higher accuracy as well as higher macro- and weighted-average F1-scores. The improvement is driven mainly by stronger classification of positive sentences.

Table E1: Positive Word Forms Included in the SHAP-Based Sentiment Dictionary

Positive Word Forms ($N = 277$)						
abate	brightest	easy	healthy	picks	settle	subdue
abated	brisk	ebb	help	poise	settled	subdued
abates	brisked	ebbed	helped	poised	settles	subdues
abating	brisking	ebbing	helpful	poises	settling	subduing
accelerate	briskly	ebbs	helpfuler	poising	sign	subside
accelerated	brisks	ebullient	helpfulest	positive	signed	subsided
accelerates	buoy	eliminate	helping	positiver	signing	subsides
accelerating	buoyant	eliminated	helps	positivest	signs	subsiding
advance	buoyanter	eliminates	impetus	progress	solid	success
advanced	buoyantest	eliminating	impetuses	progressed	solider	successes
advances	buoyed	encourage	improve	progresses	solidest	successful
advancing	buoying	encouraged	improved	progressing	spur	supportive
appear	buoys	encourages	improvement	promise	spurred	surprisingly
appeared	check	encouraging	improvements	promised	spurring	track
appearing	checked	exceptional	improves	promises	spurs	tracked
appears	checking	expand	improving	promising	stabilize	tracking
appreciable	checks	expanded	lessen	rebound	stabilized	tracks
attractive	comfortable	expanding	lessened	rebounded	stabilizes	turnaround
augur	conducive	expands	lessening	rebounding	stabilizing	underutilization
augured	confidence	fade	lessens	rebounds	stable	underutilizations
auguring	confidences	faded	lift	recede	stabler	upbeat
augurs	confident	fades	lifted	receded	stablest	uptrend
avert	consistent	fading	lifting	recedes	steadier	upturn
averted	contain	favorable	lifts	receding	steadiest	upturned
averting	contained	favorably	mitigate	recover	steadily	upturning
averts	containing	foot	mitigated	recovered	steady	upturns
beneficial	contains	footed	mitigates	recoveries	stimulate	vigor
benign	diminish	footing	mitigating	recovering	stimulated	vigorous
benigner	diminished	foots	momenta	recovers	stimulates	wane
benignest	diminishes	gain	momentum	recovery	stimulating	waned
bolster	diminishing	gained	mute	remarkable	strength	wanes
bolstered	disinflationary	gaining	muted	remarkably	strengthen	waning
bolstering	dissipate	gains	mutes	resilience	strengthened	warm
bolsters	dissipated	grew	muting	resiliences	strengthening	warmer
boost	dissipates	grow	optimism	resilient	strengthens	warmest
boosted	dissipating	growing	optimisms	robust	strengths	welcome
boosting	ease	grown	optimistic	robuster	strong	well
boosts	eased	grows	pick	robustest	stronger	
bright	eases	healthier	picked	satisfactory	strongest	
brighter	easing	healthiest	picking	satisfied	strongly	

Note: The table reports the positive word forms included in the SHAP-based sentiment dictionary. The entries are generated automatically by expanding the selected positive lemmas to their principal grammatical inflections and are listed alphabetically.

Table E2: Negative Word Forms Included in the SHAP-Based Sentiment Dictionary

Negative Word Forms ($N = 249$)						
adverse	costly	doubting	imbalance	negative	slowing	undermined
adversely	cutback	doubts	imbalances	overbuilde	slows	undermines
adverser	damage	downside	impair	overheat	sluggish	undermining
adversest	damages	downtrend	impaired	overheated	sluggisher	undershooting
backlog	decelerate	drag	impairing	overheating	sluggishest	undesirable
backlogs	decelerated	dragged	impairs	overheats	sluggishness	undesirably
bankruptcies	decelerates	dragging	impede	overstate	sluggishnesses	uneven
bankruptcy	decelerating	drags	impeded	overstated	softness	unevener
bottleneck	deceleration	drought	impedes	overstates	softnesses	unevenest
bottlenecks	decelerations	droughts	impeding	overstating	strain	unsettle
bottom	delinquencies	erode	inabilities	pessimistic	strains	unsettled
bottoms	delinquency	eroded	inability	pressure	stretch	unsettles
cause	depress	erodes	inconsistent	pressures	stretched	unsettling
caused	depressed	eroding	inhibit	problem	stretches	unsustainable
causes	depresses	erosion	inhibited	problems	stretching	unwelcome
causing	depressing	erosions	inhibiting	repercussion	subpar	vulnerabilities
challenge	deteriorate	exacerbate	inhibits	repercussions	threat	vulnerability
challenges	deteriorated	exacerbated	instabilities	reservation	threaten	vulnerable
challenging	deteriorates	exacerbates	instability	reservations	threatened	weak
complicate	deteriorating	exacerbating	insufficient	retard	threatening	weaken
complicated	deterioration	excessive	lag	retarded	threatens	weakened
complicates	deteriorations	fail	lagged	retarding	threats	weakening
complicating	difficult	failed	lagging	retards	tightness	weakens
concern	difficulties	failing	lags	risk	tightnesses	weaker
concerned	difficulty	fails	lose	risks	trouble	weakest
concerns	disappointing	falter	loses	scarcities	troubles	weakness
constrain	dislocation	faltered	losing	scarcity	turmoil	weaknesses
constrained	dislocations	faltering	loss	severely	turmoiled	worried
constraining	disrupt	falters	losses	shortage	turmoiling	worrisome
constrains	disrupted	fragile	lost	shortages	turmoils	worsen
contraction	disrupting	fragiler	mislead	shortfall	unacceptably	worsened
cool	disruption	fragilest	misleading	shortfalls	uncertain	worsening
cooler	disruptions	hard	misleads	slow	uncertainties	worsens
coolest	disrupts	hardship	misled	slowdown	uncertainty	
costlier	doubt	headwind	mismatch	slowdowns	uncomfortably	
costliest	doubted	headwinds	mismatches	slowed	undermine	

Note: The table reports the negative word forms included in the SHAP-based sentiment dictionary. The entries are generated automatically by expanding the selected negative lemmas to their principal grammatical inflections and are listed alphabetically.

Table E3: Classification Performance of the SHAP and LM Dictionaries

	SHAP Dictionary				Loughran and McDonald (2011) Dictionary				
	Precision	Recall	F1-Score	Support		Precision	Recall	F1-Score	Support
Neutral	0.58	0.47	0.52	119	Neutral	0.55	0.55	0.55	119
Positive	0.52	0.73	0.61	84	Positive	0.52	0.43	0.47	84
Negative	0.60	0.54	0.57	97	Negative	0.54	0.63	0.58	97
Accuracy			0.56	300	Accuracy			0.54	300
Macro Avg.	0.57	0.58	0.56	300	Macro Avg.	0.54	0.53	0.53	300
Weighted Avg.	0.57	0.56	0.56	300	Weighted Avg.	0.54	0.54	0.54	300

Note: The table compares the classification performance of the SHAP dictionary and the Loughran–McDonald dictionary on the same held-out evaluation sample. The sample comprises 300 manually labeled sentences that were not used to fine-tune the RoBERTa model. Precision denotes the proportion of predictions assigned to a given sentiment category that are correct, whereas recall denotes the proportion of observations in that category that are correctly identified. The F1-score is the harmonic mean of precision and recall. Support reports the number of manually labeled sentences in each category. Accuracy measures the overall proportion of correctly classified sentences. The macro average assigns equal weight to each sentiment category, whereas the weighted average accounts for differences in class size. Higher values indicate better classification performance.