Fedspeak, LLM-Derived Signals, and High-Frequency Trading

Aihui Wei*

Abstract

Financial markets are often assumed to incorporate Federal Reserve policy decisions instantaneously. Yet it remains unclear whether markets fully absorb policy-relevant information at the moment of release, or whether information continues to be revealed as the policy narrative unfolds across sequential communications. We construct a novel measure of policy communication shocks, the Communication Divergence Signal (CDS), which captures semantic, tonal, and framing shifts across sequential FOMC communications. Using large language models and high-frequency event-study methods, we show that communication divergence generates significant and persistent movements in intraday asset prices and abnormal trading volume. Abnormal trading volume is measured relative to non-announcement benchmarks at the same intraday horizon, isolating excess trading activity associated with information arrival. We further document systematic responses in daily open interest, consistent with belief updating and position reallocation rather than transitory liquidity provision. These findings highlight the role of communication sequencing and framing in shaping price discovery and trading dynamics across financial markets.

Keywords: Central Bank Communication, High-Frequency Trading, Large Language Models, Trading Volume, Monetary Policy

JEL Classifications: E52, E58, G12, C55.

*Ph.D. Program in Economics, The Graduate Center, City University of New York. Corresponding author: awei@gc.cuny.edu. I thank Sebastiano Manzan and participants at the 2026 Midwest Macroeconomics Meeting and the North American Summer Meeting of the Econometric Society for helpful comments and suggestions. All errors are my own.

1 Introduction

Central banks increasingly rely on communication as a policy instrument. Beyond setting the federal funds rate, the Federal Reserve shapes expectations through multiple communication channels. The post-meeting FOMC statement announces the policy decision, while subsequent releases, the press conference, the minutes, and policymakers' speeches, offer additional context and reveal subtle shifts in emphasis. These communications are typically longer and more complex, addressing multiple policy dimensions and reflecting collective views rather than a single decision. Understanding how markets process these layers of information is essential for assessing the effectiveness of modern monetary policy.

This paper examines how financial markets respond to new information embedded in FOMC communications released after the statement. To identify this effect, we develop measures of *communication shocks*, defined as the unanticipated divergence in meaning between the statement and subsequent channels, including the minutes, the press conference, and Fed speeches. We use a sentence-level embedding approach to capture subtle semantic shifts. These methods detect changes in tone, emphasis, and policy content that are not obvious from surface reading. By quantifying latent information beyond the initial statement that still shapes market expectations, this approach provides a new perspective on how the Federal Reserve's words move markets. Using high-frequency futures data, we then examine how tick-level federal funds futures and Treasury futures respond immediately to these shocks. The analysis looks at both price movements, which reflect changes in policy expectations, and trading volume, which reveals uncertainty and disagreement among market participants. By linking text-based measures of communication with intraday financial data, the paper bridges research on central bank communication and the literature on high-frequency event studies.

While prior research has shown that FOMC statements and minutes affect asset prices, much less is known about how markets process information across different communication channels within the same meeting. Existing studies typically analyze each document in isolation, providing limited evidence on how policy information is revealed and in-

terpreted as the Federal Reserve elaborates its message over time. This paper studies the informational gap across FOMC communication channels, focusing on how markets respond to new information released after the initial policy statement.

The paper makes three contributions. First, it introduces a language-model-based framework to measure semantic divergence across FOMC documents from the same meeting and uses LLM-derived tone scores to capture changes in the stance and emphasis of policy communication, allowing us to quantify differences in meaning and stance without relying solely on predefined keywords or dictionary-based approaches. Second, the paper extends high-frequency event study methods to precisely identify the timing and magnitude of market responses to incremental communication, isolating how new information conveyed through press conferences and minutes affects asset prices beyond what was already priced in following the policy statement. Third, the paper links semantic divergence not only to price adjustments but also to trading volume, connecting central bank communication to broader research on information revelation, uncertainty, and market microstructure.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the data sources. Section 4 outlines the textual analysis approach, Section 5 presents the methodology, and Section 7 reports the main empirical results. Section 7 provides robustness checks, and Section 8 concludes.

2 Literature Review

A large body of research has examined how to identify monetary policy shocks and trace their transmission to financial markets. Early contributions relied on structural VARs (Christiano et al., 1999), while Romer and Romer (2004) introduced a narrative approach using internal Federal Reserve forecasts to isolate exogenous shocks. Subsequent extensions include state-dependent transmission (Tenreyro and Thwaites, 2016) and forecast-error-based identification (Bachmann et al., 2022).

A related strand emphasizes the *information effect* of monetary policy announcements. Nakamura and Steinsson (2018) demonstrate that unexpected rate increases are often ac-

companied by higher output forecasts, consistent with markets learning about stronger fundamentals rather than a pure monetary tightening. Jarociński and Karadi (2020) decompose announcements into monetary policy and central bank information shocks, finding that ignoring the information component biases estimates of monetary non-neutrality. Cieslak and Schrimpf (2019) show that press conferences and minutes convey broader contextual information beyond monetary news. These studies underscore that monetary announcements reveal both policy intentions and new information about the economy, a distinction closely related to the communication shocks examined in this paper.

High-frequency event studies have become a core tool for measuring monetary policy surprises. Kuttner (2001) pioneered this approach using federal funds futures, and Gurkaynak et al. (2005) identified two distinct dimensions: a "target" factor tied to immediate rate changes and a "path" factor reflecting revisions to the expected future trajectory of policy. Subsequent refinements by Gertler and Karadi (2015), Jarociński and Karadi (2020), and Bauer and Swanson (2023) have sharpened the separation of monetary policy shocks from central bank information effects. The zero lower bound period further highlighted the role of communication: Swanson and Williams (2014) show that Treasury yields remained responsive to policy news even at the ZLB, and Campbell et al. (2012) distinguish between "Odyssean" and "Delphic" forward guidance, demonstrating that words can move markets as much as rate changes.

A growing literature examines how central bank communication shapes financial markets. Hansen and McMahon (2016) use computational linguistics to separate FOMC statements into forward guidance and current-conditions components, finding stronger market effects from forward guidance. Ehrmann and Fratzscher (2007) show that speeches and press conferences provide additional context beyond the written statement. Hansen et al. (2018) find that greater transparency has disciplined FOMC members' language, while Hubert and Labondance (2018) show that variation in tone reflects disagreement among policymakers.

Quantifying policy language is challenging, as official documents are complex and intentionally nuanced. Early approaches used word counts (Boukus and Rosenberg, 2006) and tone indices (Apel and Grimaldi, 2012), but were limited by sensitivity to specific

word choices. More recent work applies computational linguistics: Ehrmann and Talmi (2019) show that semantic novelty in wording is associated with higher market volatility, and Doh et al. (2025) fine-tune transformer models to estimate policy stance. Baumgartner and Zahner (2024) apply neural embeddings to over 23,000 central bank communications, demonstrating that embeddings capture policy signals and uncertainty. Deep learning models—including XLNet (Handlan, 2022), GPT (Hansen and Kazinnik, 2023), and BERT (Gorodnichenko et al., 2023; Curti and Kazinnik, 2023; Shah et al., 2023)—have further advanced the systematic analysis of how tone and guidance evolve across FOMC communications.

In summary, the literature has identified monetary policy shocks using a variety of methods and demonstrated that central bank communication has significant effects on financial markets, but most studies examine communication channels in isolation rather than as parts of a broader process. Recent advances in NLP make it possible to capture subtle, context-dependent changes in meaning. Building on these insights, this paper measures how the FOMC's message evolves across communication channels and examines how financial markets respond through both price movements and trading activity in a high-frequency setting.

3 Data

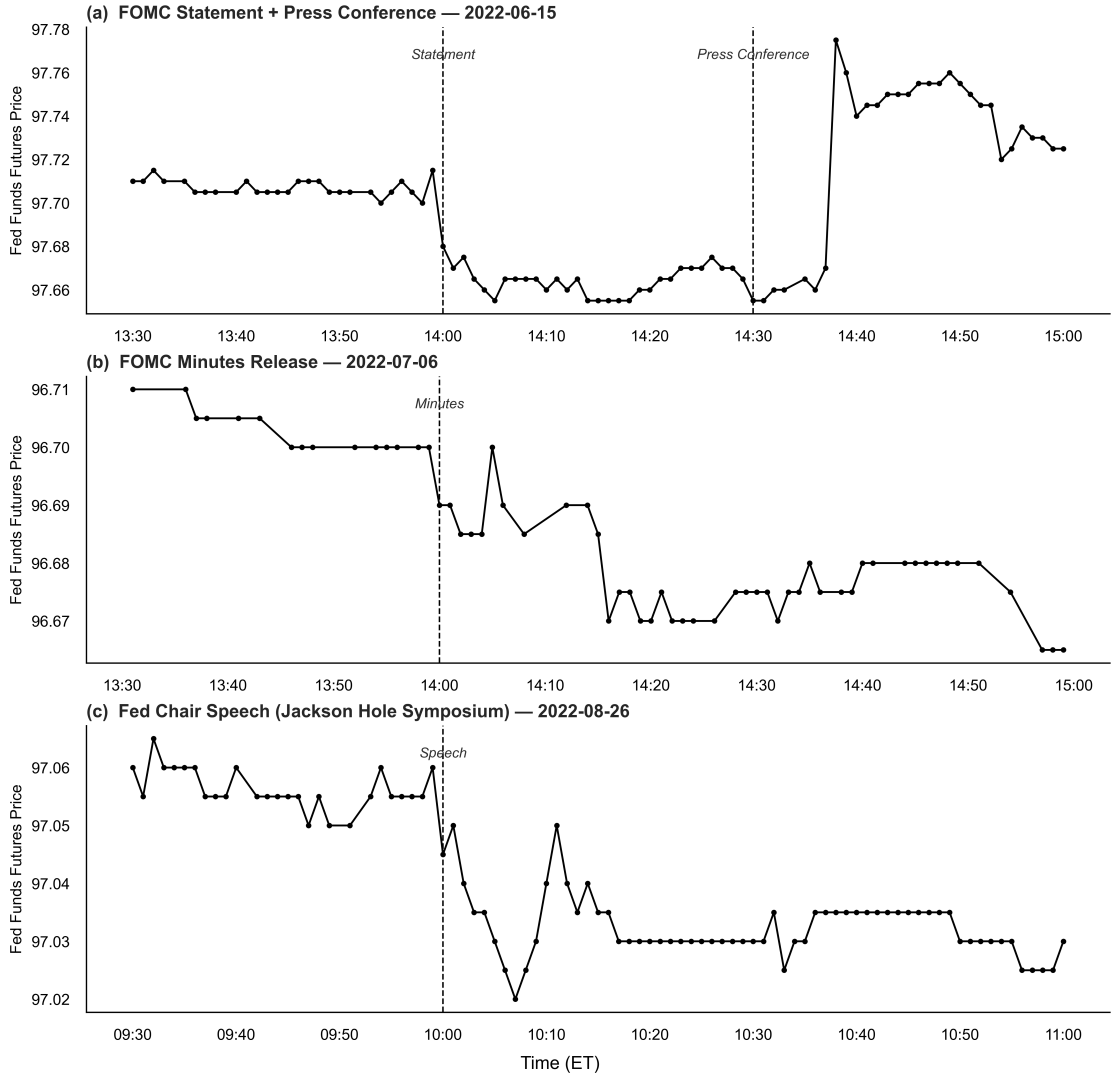
3.1 High-Frequency Trading Data

We use high-frequency trading data from several key U.S. markets to analyze asset price movements around Federal Reserve communications. The dataset spans from January 2011 to December 2024 and covers multiple asset classes: federal funds futures (ZQ), Eurodollar futures (GE), 2-year (ZT), 5-year (ZF), and 10-year (ZN) Treasury note futures, E-mini S&P 500 futures (ES), and EUR/USD currency futures (6E). All contracts are sourced from the CME Group and include tick-level and minute-level observations for precise tracking of intraday market reactions. Detailed contract specifications are provided in Appendix A.

The sample data covers the period from 2010 to 2025 and includes all scheduled FOMC

meetings during this period. We process large tick-level datasets across multiple assets and event windows in parallel using multi-core computing, which substantially improved efficiency while keeping results identical. Each FOMC communication time is converted to UTC to maintain consistency across sources. Figure 1 illustrates how federal funds futures prices move around key FOMC communication events, including policy announcements, minutes releases, and speeches (see Appendix B for detailed narratives of each episode).

Figure 1: Intraday Reactions of Federal Funds Futures Prices



Notes: The figure illustrates intraday adjustments in federal funds futures prices surrounding major FOMC communication events in 2022. Each series corresponds to the most actively traded (front-month) contract, and we sampled the tick-level federal funds futures data at one-minute frequency.

3.2 Daily Financial Market Data

To complement the high-frequency identification strategy, we also use daily financial market data. Daily observations allow us to examine longer-run market adjustments across different monetary policy regimes, including the zero lower bound and the post-pandemic tightening cycle. The daily dataset spans January 1990 to January 2025 and covers major U.S. asset classes. We use the S&P 500 index to capture broad equity market performance, along with two-year, five-year, and ten-year Treasury yields from the Federal Reserve H.15 release. We also include settlement prices for federal funds futures from CME and Bloomberg, which provide a daily measure of changes in expected monetary policy. These series offer long historical coverage and help assess the persistence of market responses to Federal Reserve communication.

In addition, we collect daily open interest for major U.S. futures contracts. Open interest reflects market participation and positioning around policy events. It helps identify whether traders adjust exposures following Federal Reserve announcements. Daily contract-level open interest series are obtained from the CME Group historical portal[1]. Further details and data checks are provided in Appendix H.

3.3 Textual Data

We draw on four forms of FOMC communication: the post-meeting statement, the minutes, the press conference, and speeches by Federal Reserve officials. All documents are collected from the Federal Reserve's official website using web-scraping techniques and standardized for consistency. The sample spans 2010–2025 (when speech timestamps become consistently available) and includes 120 policy statements, 115 minutes documents, 95 press conference transcripts, and more than 900 speeches. Unscheduled or emergency meetings are excluded.

Statements are released immediately after each meeting and anchor market expectations. Press conferences (available from 2011; after every meeting since 2019) include the Chair's prepared remarks and a Q&A session. Minutes are published three weeks af-

[1]Accessed via an automated Python script. Source: CME Group, "Equity Index Products Volume and Open Interest" and "Interest Rate Products Volume and Open Interest", https://www.cmegroup.com/.

ter each meeting; to isolate new information, we extract only the substantive discussion sections (e.g., "Staff Economic Outlook," "Participants' Views," and "Committee Policy Action") and remove boilerplate language. Speeches elaborate on policy themes outside the formal post-meeting cycle and are delivered by the Chair, Vice Chair, Governors, and Reserve Bank Presidents at various venues.

Precise event timing is essential for matching Federal Reserve communications to high-frequency market data. I construct a consistent timestamp dataset for all FOMC communication channels, converting all times to UTC. Statements, minutes, and press conferences follow established release schedules, while speech timestamps are extracted from official Federal Reserve documents (see Appendix E for the full timestamp protocol and Appendix Table E2 for the release schedule).

4 Textual Analysis of FOMC Communications

Standard text preprocessing is applied to all FOMC documents, including stop-word removal, tokenization, negation handling using the Loughran and McDonald (2011) sentiment lexicon, lemmatization, and standardization of economic acronyms. Details are provided in Appendix D.

4.1 LLM-Based Tone Scoring

We use a large language model as a measurement technology to map qualitative policy language into a continuous tone score. For each Federal Reserve communication document i, the model produces a scalar tone score $T_i \in [-1, 1]$. Higher values indicate a more positive or optimistic evaluative tone. Lower values indicate a more negative or cautious tone.

The tone score is intended to capture the general, policy-relevant sentiment conveyed by the communication. The model is instructed to rely only on the language contained in the document itself. It abstracts from reported market reactions, asset price movements, and realized economic outcomes. As a result, the tone score reflects the ex ante evaluative content of the communication at the time of release.

We use ChatGPT-4o to implement this measurement. The model evaluates the overall tone conveyed by each document based on contextual language use. This approach accounts for negation, modality, and forward-looking statements. These features are important in central bank communication and are difficult to capture using dictionary-based methods.

When multiple documents belong to the same communication channel within meeting t, we aggregate tone scores at the channel level. Specifically, the channel-level tone score is defined as the average across documents:

$$T_t^{(j)} = \frac{1}{N_{j,t}} \sum_{i \in \mathcal{D}_{j,t}} T_i, \tag{1}$$

where $\mathcal{D}_{j,t}$ denotes the set of documents in channel $j \in \{\text{minutes}, \text{press}, \text{speech}\}$ associated with meeting t. The policy statement released at meeting t yields the corresponding tone score $T_t^{(\text{statement})}$. The model is prompted to evaluate overall tone based on assessments of economic conditions, risks, and the outlook, producing a continuous score from -1 (strongly negative) to $+1$ (strongly positive). The full prompt specification is provided in Appendix **??**.

4.2 Tone Gap

To quantify informational shifts between the initial policy statement and subsequent Federal Reserve communications, we define the *tone gap*, which measures changes in the policy-relevant tone conveyed across communication channels following each FOMC meeting. Let $\mathcal{P}_t = \{\text{minutes, press, speech}\}$ denote the set of post-statement communications associated with meeting t. For each channel $j \in \mathcal{P}_t$, the tone gap is defined as:

$$\text{ToneGap}_t^{(j)} = T_t^{(j)} - T_t^{(\text{statement})}, \tag{2}$$

where $T_t^{(\text{statement})}$ denotes the LLM-based tone score of the policy statement released at meeting t, and $T_t^{(j)}$ is the corresponding tone score for post-statement communication channel j (i.e., meeting minutes, press conference transcript, or Chair speeches).

A positive value of $\text{ToneGap}_t^{(j)}$ indicates that the subsequent communication adopts a more hawkish policy interpretation relative to the initial statement, while a negative value reflects a shift toward a more dovish interpretation. The tone gap therefore captures both the direction and magnitude of changes in the Federal Reserve's policy narrative as it unfolds across communication formats following each meeting.

To summarize the overall evolution of tone across post-meeting communications, we construct an *aggregate tone gap* for each meeting as a weighted average of the channel-specific tone gaps:

$$\overline{\text{ToneGap}}_t = \sum_{j \in \mathcal{P}_t} w_j \, \text{ToneGap}_t^{(j)}, \qquad \sum_{j \in \mathcal{P}_t} w_j = 1, \tag{3}$$

where w_j denotes the weight assigned to communication channel j. In the baseline specification, we assign equal weights across channels. The aggregate tone gap, $\overline{\text{ToneGap}}_t$, captures the net shift in policy-relevant tone across the entire post-meeting communication cycle. Positive (negative) values indicate that, on average, subsequent communications convey a more hawkish (more dovish) interpretation relative to the initial policy statement.

4.3 Semantic Divergence Measurement

We use the `LLM2Vec` (Chen et al., 2024) framework as the primary embedding model to measure semantic divergence.[2] The model generates context-aware vector representations that integrate information from multiple transformer layers. This design allows the embeddings to reflect both local sentence meaning and broader document-level structure.

Compared with standard sentence-transformer models such as SBERT, `LLM2Vec` better preserves long range semantic relationships and overall textual coherence. This feature is particularly important for FOMC communications, where changes in meaning often arise through subtle rephrasing rather than explicit shifts in language. Each document is represented by a dense embedding vector that summarizes its semantic content. We compute the cosine similarity between the policy statement and the corresponding FOMC

[2]See Appendix I.2 for details on the training procedure.

communication to obtain a continuous measure of semantic divergence. Lower similarity values indicate greater informational distance between the two texts.

Importantly, the semantic similarity measure underlying our divergence metric is designed to capture novelty in propositional content across communications. It does not reflect differences in evaluative tone, which are measured separately using the sentiment-based approach described elsewhere in the paper.

For robustness, we also replicate the analysis using the `all-MiniLM-L6-v2` model from the Sentence-BERT (SBERT) family, a common benchmark for semantic similarity. Results remain qualitatively consistent, indicating that the observed communication gaps are not specific to the choice of embedding model.

All models are implemented in PyTorch using the Hugging Face Transformers framework. Training details, text cleaning procedures, and model specifications are provided in Appendix I.

4.3.1 Embedding Generation

After preprocessing, we generate document-level embeddings using the `LLM2Vec` framework. The policy statement and the extracted portions of each post-meeting communication are encoded into dense vector representations that summarize their semantic content. Unlike standard sentence-transformer models, `LLM2Vec` adapts a large language model into a bidirectional text encoder through unsupervised post-training. This approach allows the embeddings to incorporate both local sentence-level cues and broader document-level context.

As a result, The embedding process compresses the full text of each document into a fixed-length vector while preserving meaningful variation in content. This design is well suited for FOMC communications, where new information is often conveyed through shifts in narrative structure, conditional language, or the relative emphasis placed on policy considerations, rather than through explicit changes in wording.

4.3.2 Semantic Divergence Computation

I measure how similar the policy statement is to each post-meeting communication channel using cosine similarity:

$$\text{Sim}_t^{(j)} = \frac{\mathbf{v}_t^{\text{statement}} \cdot \mathbf{v}_t^{(j)}}{\|\mathbf{v}_t^{\text{statement}}\| \cdot \|\mathbf{v}_t^{(j)}\|}, \qquad j \in \mathcal{P}_t, \tag{4}$$

Here, $\mathbf{v}_t^{\text{statement}}$ and $\mathbf{v}_t^{(j)}$ are the LLM2Vec embeddings of the statement and post-meeting channel j (such as the minutes, press conference, or Fed Chair's speech) for meeting t. The similarity score, $\text{Sim}_t^{(j)}$, ranges from -1 to 1. Higher values indicate stronger correspondence in content and emphasis. Lower values signal greater semantic divergence. Then I define semantic divergence as the complement of similarity:

$$\text{SemDiv}_t^{(j)} = 1 - \text{Sim}_t^{(j)}, \qquad j \in \mathcal{P}_t. \tag{5}$$

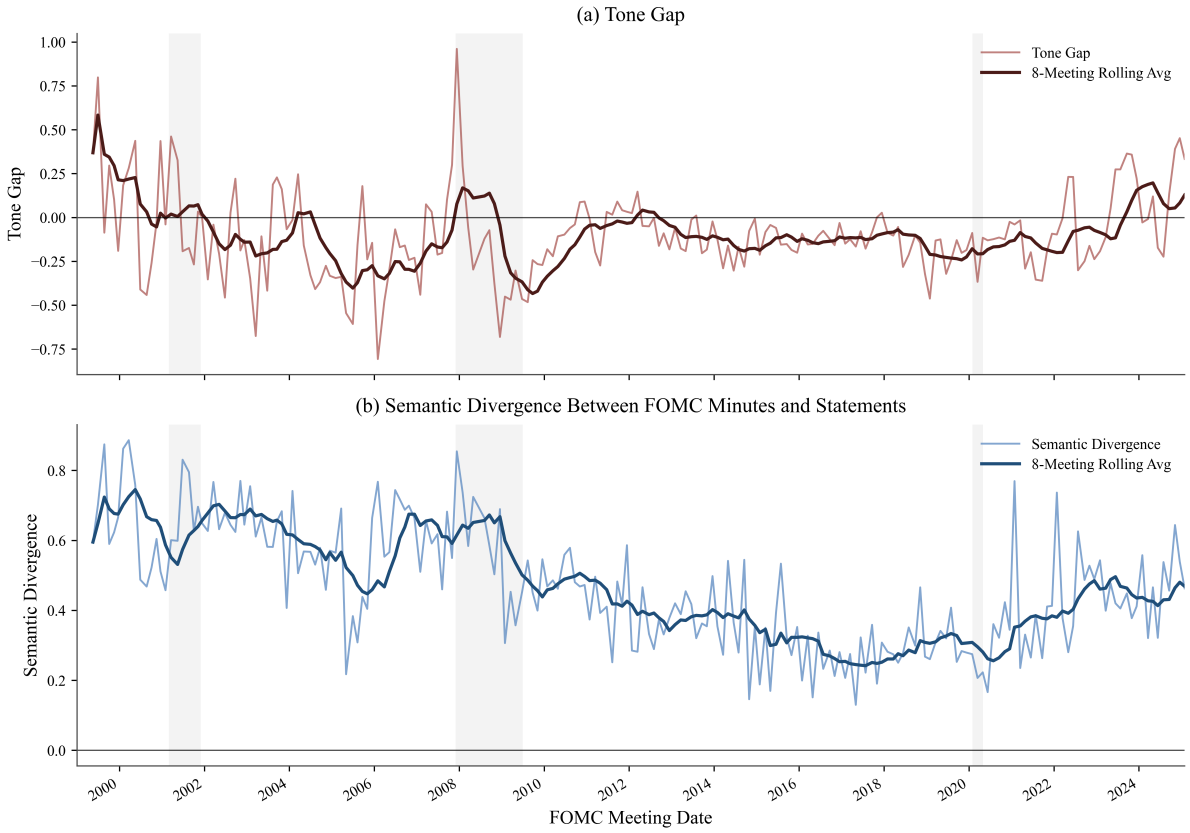
High similarity signals repetition, while high divergence points to new policy information. At the December 2007 FOMC meeting, the statement described "unusually uncertain" financial conditions and announced a 25 basis points (bp) rate cut. The minutes that followed expressed stronger concern and noted that some members supported a larger cut. They also introduced the Term Auction Facility[3], a liquidity tool not mentioned in the statement. Although the overall tone differed only slightly, the cosine similarity between the two documents declined sharply, producing a local trough in Figure 2. This divergence captures a shift in emphasis rather than sentiment, as the minutes disclosed new information about risks and policy deliberations that the statement withheld.

This case shows how semantic divergence complements sentiment analysis. The similarity measure captures how much additional context the minutes introduce, while the tone gap reflects shifts in sentiment. Together, they describe how FOMC communication evolves between the statement and the minutes, offering insight into how new details

[3]The Term Auction Facility (TAF) was a temporary program under which the Federal Reserve auctioned term funds to depository institutions to ease pressures in short-term funding markets. The final TAF auction was conducted on March 8, 2010. See *Federal Reserve Board – Term Auction Facility*, https://www.federalreserve.gov/monetarypolicy/taf.htm.

and emphasis shape market perception.

Figure 2: FOMC Statements vs. Minutes



Notes: Thin lines show raw series; thick lines show 8-meeting rolling averages. Panel (a) plots the tone gap, Panel (b) shows semantic divergence. Shaded areas denote NBER recessions.

4.4 Reduced-Form Belief Updating Framework

Let $\mu_t \equiv \mathbb{E}_t[f(\theta_t)]$ denote market expectations about a policy-relevant latent state θ_t, such as the expected future path of the federal funds rate. Following the information-based monetary policy literature (e.g., Nakamura and Steinsson (2018); Swanson (2021)), we interpret μ_t as the latent object that asset prices respond to contemporaneously when new policy-relevant information is revealed through central bank communication.

Subsequent post-meeting communications provide additional public information that

may revise these beliefs. We model belief updating in reduced form as

$$\mu_t^{(j)} = \mu_t + \Delta\mu_t^{(j)}, \tag{6}$$

where $\mu_t^{(j)}$ denotes market beliefs after incorporating information from post-meeting communication channel j, and $\Delta\mu_t^{(j)}$ captures the corresponding informational innovation.

We assume that the magnitude of belief revision depends on two components: (i) the intensity of evaluative tone conveyed in the communication, and (ii) the novelty of its propositional content relative to the policy statement. Accordingly, we write

$$|\Delta\mu_t^{(j)}| \propto |T_t^{(j)} - T_t^{\text{statement}}| \times \text{SemDiv}_t^{(j)}, \tag{7}$$

where $T_t^{(j)}$ denotes the LLM-based tone score of post-meeting communication j, $T_t^{\text{statement}}$ is the tone of the policy statement, and $\text{SemDiv}_t^{(j)}$ measures semantic divergence from the statement.

This reduced-form relationship provides a direct link between belief updating and observable features of policy communication. In the empirical analysis, we summarize the magnitude of belief revision using a text-based communication signal that combines tonal shifts and semantic novelty across post-meeting communications. In the next subsection, we formally define this Communication Divergence Signal (CDS) and discuss its construction.

4.5 Communication Divergence Signal

To summarize the informational content of post-meeting communications relative to the policy statement, we construct a Communication Divergence Signal (CDS), which serves as a reduced-form measure of qualitative policy communication.

Formally, for meeting t, we define

$$CDS_t \;\equiv\; \text{ToneGap}_t \times \text{SemDiv}_t \;=\; \text{ToneGap}_t \times (1 - \text{Sim}_t). \tag{8}$$

The CDS increases when post-meeting communications both (i) shift in tone relative

to the policy statement and (ii) introduce substantively new policy-relevant content. By construction, the CDS captures the joint intensity and novelty of qualitative policy communication.

5 Methodology

5.1 High-Frequency Market-Based Surprises

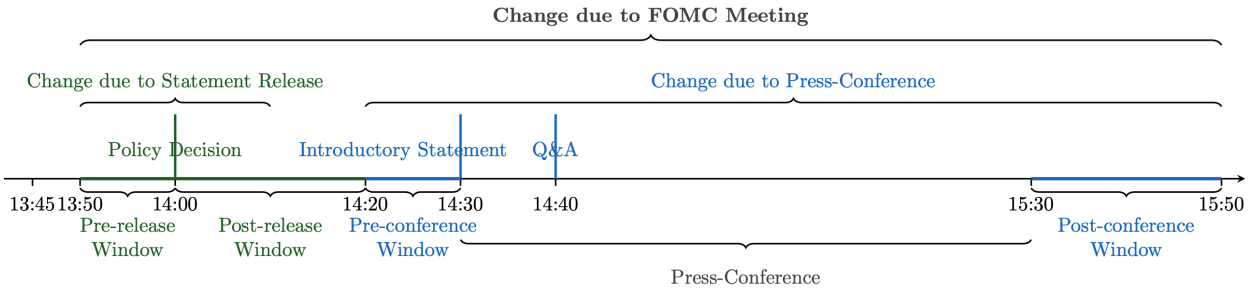
To identify monetary policy surprises at high frequency, I use transaction-level Federal Funds Futures (FFF) data from the CME Group. Let t_0 denote the time of the FOMC announcement (typically 2:00 p.m. ET). Following Gürkaynak et al. (2005, henceforth GSS), I examine asset price movements within a narrow window surrounding each communication:

$$t_- = t_0 - 10 \text{ minutes}, \qquad t_+ = t_0 + 20 \text{ minutes}.$$

This window isolates the immediate market response to monetary policy news while limiting the influence of subsequent trading and other macroeconomic factors.

As illustrated in Figure 3, each FOMC announcement comprises several communication components, including the policy decision, introductory statement, and the Chair's press conference.

Figure 3: Timeline of Intraday FOMC Policy Communications



Notes: The figure shows the sequence of intraday communication events during a typical FOMC meeting day. The policy statement is released at 2:00 p.m. Eastern Time, followed by the Chair's press conference at 2:30 p.m. Brackets indicate the pre-release, post-release, and post-conference windows used for the high-frequency analysis.

Among all FFF contracts traded during this window, the most actively traded contract c^* is selected as

$$c^* = \arg\max_{c \in \mathcal{C}} \text{Volume}_c,$$

so that the estimated surprise reflects the most liquid and information-relevant segment of the market. Let P_{pre} and P_{post} denote the last transaction prices before and after the window, respectively, for contract c^*. The corresponding monetary policy surprise, expressed in basis points, is defined as

$$\Delta i_t = -100\,(P_{\text{post}} - P_{\text{pre}}). \tag{9}$$

The negative sign converts futures price changes into rate changes, so that an increase in futures prices implies a decline in the expected federal funds rate. Equation (9) therefore captures the unanticipated component of the policy decision as inferred from market-implied expectations at the time of each FOMC communication.

For each asset, I identify the contract with the highest tick volume within a symmetric window centered around the policy release time:

$$[t_0 - W,\, t_0 + W], \qquad W \in \{10, 30, 60, 90, 120\} \text{ minutes}.$$

This liquidity-based selection ensures that price reactions are derived from the most actively traded contract during the announcement period, capturing the market segment where information is incorporated most efficiently.

For every timestamp t, the midpoint price is computed as

$$\text{Mid}_t = \frac{\text{High}_t + \text{Low}_t}{2},$$

which provides a smoothed measure of transaction value that mitigates the influence of bid–ask spreads and other sources of microstructure noise.

In contrast to Δi_t, which measures the unanticipated component of the policy decision inferred from federal funds futures, $\Delta P_{i,t}$ captures the market price response of asset i to

information released around the FOMC communication.

Let t_0 denote the announcement time of the FOMC communication. Let $P_i(t)$ denote the midpoint price of asset i at time t. The price change of asset i over a symmetric event window is defined as

$$\Delta P_{i,t}(W) = P_i(t_0 + W) - P_i(t_0 - 10),$$

where $W \in \{10, 30, 60, 90, 120\}$ minutes. This measure captures the net price adjustment of asset i over the event window surrounding the FOMC announcement.

6 Identification of Text-Based Communication Shocks

This section describes how we separate anticipated and unanticipated components of post-meeting policy communication. Our objective is to isolate the portion of textual information that is new to markets at the time of release.

We represent market anticipation of policy communication using a proportional decomposition of the realized textual signal:

$$\mathbb{E}[CDS_t \mid \Omega_{t-}] = p_t \cdot CDS_t, \quad p_t \in [0, 1]. \tag{10}$$

In a reduced-form sense, The parameter p_t summarizes the extent to which the content of a given communication appears to have been anticipated and absorbed by markets prior to its release. Values of p_t close to one indicate that a large fraction of the signed communication signal was effectively priced in, while values close to zero indicate largely unanticipated releases.

This proportional representation should be interpreted as a normalization that decomposes the realized communication signal into an anticipated component and an unexpected component. It is not intended as a structural model in which markets literally observe or form expectations over CDS_t prior to release. This specification deliberately avoids imposing a structural model of expectation formation. Instead, it provides a parsimonious reduced-form characterization of anticipation in a high-dimensional textual environment.

Using (10), we define the communication shock as the expectation innovation:

$$\text{CommShock}_t = CDS_t - \mathbb{E}[CDS_t \mid \Omega_{t-}] = (1 - p_t)\, CDS_t. \tag{11}$$

This object captures the signed component of communication that is new to markets at the time of release. Under standard high-frequency identification assumptions, only this unanticipated component should affect asset prices within a narrow announcement window.

6.1 Identification Strategy

The main empirical challenge is that the market's expectation of communication, $\mathbb{E}[CDS_t \mid \Omega_{t-}]$, is not observed. Rather than imposing a parametric model of expectation formation, we discipline anticipation using high-frequency market responses to policy communication. Intuitively, if a communication conveys more new information to markets, it should trigger a larger repricing response within a narrow announcement window on average.

Let FFF_t denote the change in federal funds futures prices measured in a narrow window around the release of post-meeting communication. A large literature shows that such high-frequency futures movements provide a clean and widely used measure of monetary policy surprises (e.g., Gürkaynak et al. (2005), Nakamura and Steinsson (2018), Swanson (2021)). Within a short announcement window, price changes primarily reflect newly revealed policy-relevant information, with limited contamination from other shocks. Because FFF_t is measured around the communication release while holding the policy decision fixed, it captures the incremental repricing induced by the communication itself.

Our identification strategy relies on a weak monotonicity assumption about how futures markets aggregate policy-relevant information. Communications that convey more new information tend to induce larger price adjustments in futures markets on average. These responses may be noisy due to liquidity conditions or market microstructure frictions, but their relative ordering remains informative.

Assumption 1 (Monotonic response) *Within a narrow announcement window, intraday fu-*

tures surprises respond monotonically, in expectation, to the unexpected component of policy communication.

Under this assumption, intraday futures surprises satisfy

$$FFF_t = g(\text{CommShock}_t) + \varepsilon_t, \tag{12}$$

where $g(\cdot)$ is a strictly increasing function and ε_t captures transitory noise and microstructure frictions. This assumption implies an ordinal restriction that does not depend on the functional form of $g(\cdot)$. This ordinal restriction implies that high-frequency futures surprises provide a noisy but informative signal about the relative ordering of unexpected communication shocks across meetings.

Building on this insight, we adopt a rank-based identification strategy that exploits ordinal information rather than the magnitude of price changes. Our approach follows the rank-based framework developed in Doh et al. (2026), which infers market expectations by maximizing the rank correlation between high-frequency asset price reactions and the unexpected component of policy communication. We identify the anticipation parameters $\{p_t\}$ by aligning the ordering of implied communication shocks with the ordering of futures surprises across meetings.

Using the structural representation $CommShock_t = (1 - p_t)CDS_t$, we therefore choose the anticipation parameters $\{p_t\}$ to maximize the number of pairwise comparisons implied by this ordinal restriction:

$$\max_{\{p_t\}} \sum_{t \neq s} \mathbf{1}(FFF_t > FFF_s)\, \mathbf{1}((1 - p_t)CDS_t > (1 - p_s)CDS_s)\,, \quad \text{s.t. } p_t \in [0, 1]\ \forall t. \tag{13}$$

The objective in (13) counts the fraction of *concordant* pairs. A pair (t, s) is concordant if the meeting associated with the larger futures surprise is also assigned the larger implied communication shock. Pairs that violate the ordering are *discordant* and contribute zero. Importantly, this criterion uses only ordinal information. It is invariant to monotone transformations of FFF_t and to any rescaling of CDS_t, and therefore robust to outliers, nonlinear price responses, and measurement units.

Equivalent rank-loss form. It is convenient to rewrite (13) using a compact notation. Define

$$S_t \equiv \text{CommShock}_t = (1 - p_t)CDS_t,$$

and the pairwise ordering indicator $y_{ts} \equiv \text{sign}(FFF_t - FFF_s) \in \{-1, +1\}$ for $t \neq s$. Maximizing the number of concordant pairs is then equivalent to minimizing the number of mis-ordered pairs:

$$\min_{\{p_t\}} \sum_{t \neq s} \mathbf{1}\big(y_{ts}\left(S_t - S_s\right) \leq 0\big), \quad \text{s.t. } p_t \in [0, 1]\ \forall t. \tag{14}$$

The objective in (14) penalizes pairwise violations of the ordering implied by futures surprises. A pair (t, s) contributes one unit of loss if the relative ordering of the implied communication shocks conflicts with the relative ordering of FFF_t and FFF_s, and zero otherwise. Equivalently, (14) minimizes the number of *discordant* pairs, or maximizes Kendall's τ between S_t and FFF_t.

Importantly, this formulation relies exclusively on ordinal information. Identification is driven by relative comparisons across meetings rather than by the magnitude of price responses. By construction, the criterion is invariant to strictly monotone transformations of FFF_t and to rescaling of CDS_t, and is therefore robust to outliers, nonlinear price reactions, and measurement units.

Final communication shock measure. Given the estimated anticipation parameters $\hat{p}_t$ obtained from the rank-based calibration, we construct the operational communication shock as

$$\widehat{\text{CommShock}}_t = (1 - \hat{p}_t)\, CDS_t. \tag{15}$$

This measure isolates the unanticipated component of post-meeting policy communication by down-weighting textual signals that were largely priced in prior to the announcement. As such, it captures variation in communication that is new to market participants and serves as the main text-based shock in the event-study and regression analyses that follow.

6.2 Empirical Specification

This subsection examines how financial markets respond to FOMC communication shocks. For each asset i and meeting t, I estimate the high-frequency response of asset prices to unexpected information contained in post-meeting communications. Given the narrow announcement window, additional controls are not included, as other macroeconomic news is unlikely to vary systematically within this period.

The baseline specification is

$$\Delta P_{i,t} = \alpha_i + \beta_i \, \text{CommShock}_t + \varepsilon_{i,t}, \tag{16}$$

where $\Delta P_{i,t}$ denotes the intraday price change constructed in Section 5.1, and CommShock_t measures the unanticipated informational component of post-meeting communication relative to the initial policy statement. The coefficient β_i captures the contemporaneous sensitivity of asset i to unexpected changes in the qualitative content of FOMC communications.

Equation (16) is estimated by ordinary least squares with heteroskedasticity-robust standard errors clustered at the meeting level. Clustering by meeting accounts for correlated residuals across assets that respond to the same communication event. The communication shock is standardized to have zero mean and unit variance, so that β_i can be interpreted as the average change in asset i's price, in basis points, for a one standard deviation increase in the communication shock.

To examine heterogeneity in responses across asset classes, I also estimate pooled specifications within broad asset categories. Assets are grouped into bonds, equities, foreign exchange, and commodities. For each class c, I estimate

$$\Delta P_{i,t} = \alpha_i + \beta^{(c)} \, \text{CommShock}_t + \varepsilon_{i,t}, \quad i \in c, \tag{17}$$

where α_i denotes asset fixed effects and $\beta^{(c)}$ summarizes the average sensitivity of asset prices within class c to communication shocks. This specification ensures that estimated responses are not driven by offsetting movements across asset classes with opposing di-

rectional reactions, and instead captures within-class transmission of qualitative policy information.

Finally, to assess the aggregate market response, I estimate a pooled specification that combines all assets:

$$\Delta P_{i,t} = \delta_i + \beta \, \text{CommShock}_t + \varepsilon_{i,t}, \tag{18}$$

where δ_i denotes asset fixed effects and β captures the average market-wide sensitivity to communication shocks. This pooled regression summarizes the overall impact of post-meeting communication on financial conditions across markets. Together, these specifications describe how qualitative information released after FOMC meetings is absorbed into asset prices and transmitted through financial markets in real time.

7 Empirical Results

In this section, I examine how tone and content evolve jointly across FOMC communications. It introduces the Communication Divergence Signal (CDS), which combines differences in tone with measures of semantic novelty. The index captures both the direction and informational content of policy language, providing a framework to study how new narratives are incorporated into financial markets.

Figure 4 below shows how tone and content evolve across FOMC communications. Panel (a) illustrates that tone differences between the minutes and the policy statement become larger when the two documents differ more in meaning. This pattern suggests that when the minutes introduce new information or highlight new risks, they also tend to shift the overall tone of communication. Panel (b) presents the Communication Divergence Signal (CDS). It captures how much the tone and content of the minutes differ from the statement released after the same meeting. High values of the CDS mark episodes when the Committee's message changes both in direction and in substance. Large movements in the CDS appear during key periods of economic uncertainty, including the early 2000s, the global financial crisis, and the pandemic. These peaks suggest that the FOMC adjusted its post-meeting narrative more sharply in times of stress or disagreement. Overall, the CDS summarize communication shocks that combine shifts in tone with new pol-

icy information.

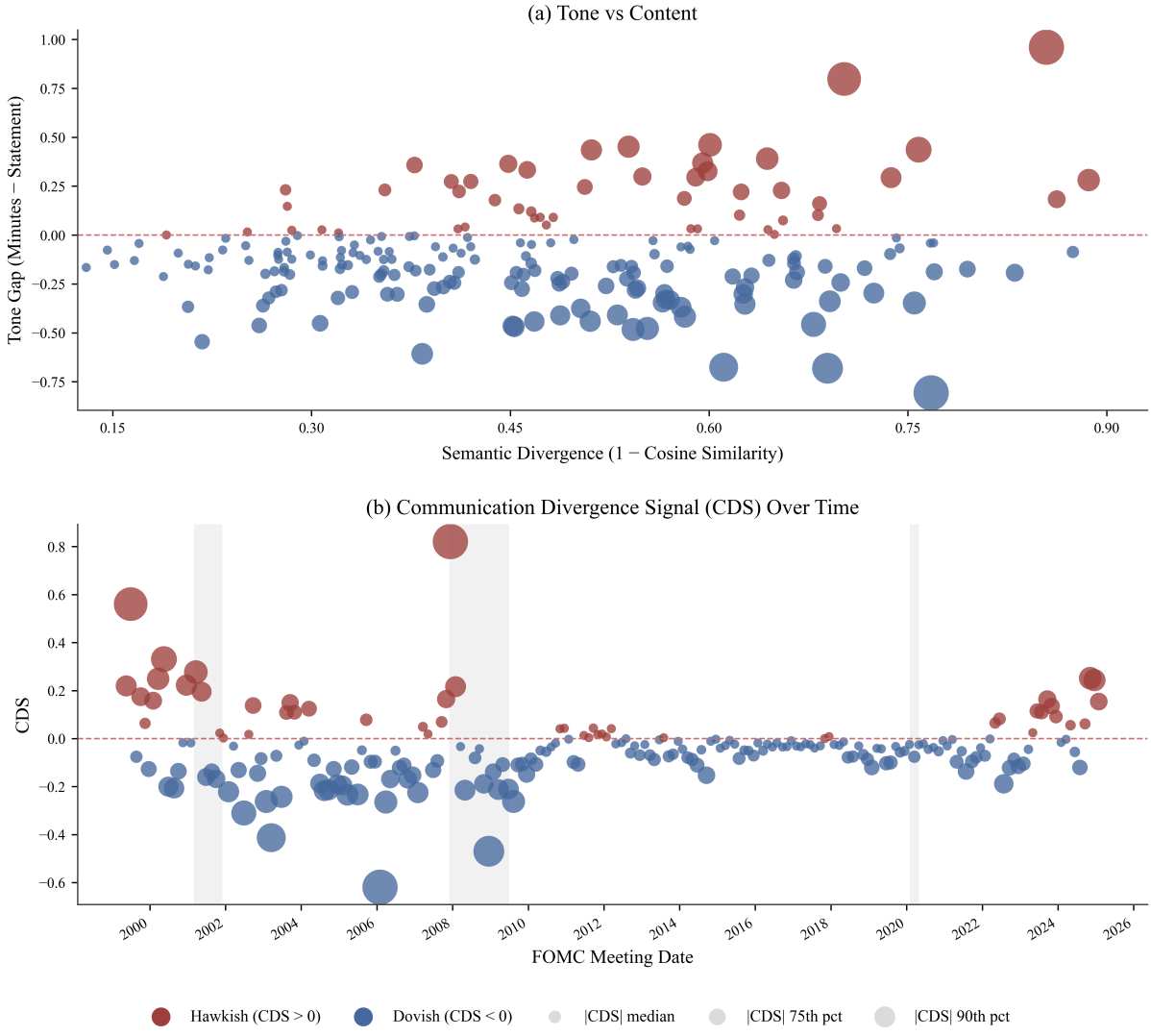
7.1 High-Frequency Price Responses

Table 1 reports how asset prices respond to the *Statement–Minutes Communication Shock* across intraday windows from 5 to 90 minutes following the release of the FOMC minutes. Equity markets show a strong and negative response. For S&P 500 futures, the estimated coefficients range from –131 to –198. A one-standard-deviation increase in the *Statement–Minutes Communication Shock* leads to a decline of roughly 1–3 percent, or about 100–300 basis points, within the first thirty minutes after the release. Since the dependent variable measures changes in futures prices in index points, these estimates reflect realistic short-term adjustments in equity markets. The magnitude is consistent with prior intraday studies, which find that equity prices typically drop between 0.5 and 2 percent following hawkish policy signals (Lucca and Moench, 2015; Handlan, 2022).

Treasury futures move in the opposite direction, showing positive and highly significant coefficients. The response becomes stronger for longer maturities, increasing from about 1.3 for the 2-year contract to more than 10 for the 10-year contract after ninety minutes. This pattern suggests that the communication shock influences both near-term policy expectations and longer-term yield dynamics, as markets reassess the anticipated path of monetary tightening.

Foreign exchange markets also react to the tone divergence between statements and minutes. Euro FX futures display consistently negative coefficients between -0.07 and -0.11 across all windows, implying a stronger U.S. dollar after relatively hawkish minutes. The cross-asset consistency supports the interpretation that markets view the tone difference between statements and minutes as a source of meaningful policy information. Overall, prices adjust most within the first thirty minutes and remain statistically significant up to ninety minutes.

Figure 4: Tone and Communication Divergence



Notes: Panel (a) plots the relationship between tone gap and semantic divergence across FOMC statements and minutes. Panel (b) shows the time series of the Communication Divergence Signal (CDS). Bubble size represents the absolute magnitude of CDS, with red (blue) indicating hawkish (dovish) post-meeting communication. Shaded areas denote NBER recession periods.

Table 1: Trading Price Response to Statement–Minutes Communication Shock

	S&P 500 Futures	2-Year Treasury Futures	5-Year Treasury Futures	10-Year Treasury Futures	Euro FX Futures
5-minute window					
CommShock	-130.78^{**}	1.35^{***}	3.19^{***}	6.57^{***}	-0.07^{***}
	(55.44)	(0.28)	(0.76)	(0.99)	(0.01)
R^2	0.09	0.30	0.32	0.33	0.12
10-minute window					
CommShock	-165.10^{***}	1.71^{***}	3.85^{***}	7.38^{***}	-0.08^{***}
	(42.08)	(0.36)	(0.86)	(1.06)	(0.01)
R^2	0.09	0.33	0.34	0.35	0.14
30-minute window					
CommShock	-198.17^{***}	2.11^{***}	4.55^{***}	8.97^{***}	-0.11^{***}
	(61.85)	(0.45)	(0.91)	(1.28)	(0.02)
R^2	0.04	0.37	0.38	0.37	0.19
60-minute window					
CommShock	-140.40^{**}	2.25^{***}	5.09^{***}	9.44^{***}	-0.11^{***}
	(71.45)	(0.38)	(1.01)	(1.53)	(0.02)
R^2	0.02	0.36	0.39	0.38	0.19
90-minute window					
CommShock	-139.30	2.26^{***}	5.63^{***}	10.20^{***}	-0.10^{***}
	(118.16)	(0.37)	(1.08)	(1.56)	(0.02)
R^2	0.01	0.34	0.39	0.39	0.13
Observations: 118					

Notes: Standard errors are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 2: Trading Price to Statement–Press Conference Communication Shock

	S&P 500	2-Year Treasury	5-Year Treasury	10-Year Treasury	Euro FX
5-min Window					
Coefficient	-64.41	0.46*	1.09*	1.29	-0.03*
Std. Error	(108.44)	(0.28)	(0.63)	(0.85)	(0.02)
p-value	[0.553]	[0.102]	[0.084]	[0.136]	[0.079]
R^2	0.01	0.03	0.02	0.02	0.03
10-min Window					
Coefficient	-21.26	0.68*	1.74*	1.89*	-0.02
Std. Error	(121.42)	(0.39)	(0.92)	(1.09)	(0.02)
p-value	[0.861]	[0.079]	[0.057]	[0.082]	[0.334]
R^2	0.00	0.02	0.03	0.02	0.01
30-min Window					
Coefficient	-209.38	2.22***	5.24**	6.67**	-0.08**
Std. Error	(243.98)	(0.82)	(2.14)	(2.81)	(0.04)
p-value	[0.391]	[0.007]	[0.014]	[0.018]	[0.039]
R^2	0.01	0.08	0.08	0.08	0.06
60-min Window					
Coefficient	-181.07	2.69***	5.69***	7.06**	-0.09*
Std. Error	(222.66)	(0.78)	(2.22)	(3.11)	(0.05)
p-value	[0.416]	[0.001]	[0.010]	[0.023]	[0.077]
R^2	0.01	0.10	0.08	0.07	0.06
90-min Window					
Coefficient	-264.23	2.55***	5.42**	6.83*	-0.08
Std. Error	(289.16)	(0.90)	(2.46)	(3.57)	(0.05)
p-value	[0.361]	[0.006]	[0.028]	[0.056]	[0.115]
R^2	0.01	0.08	0.06	0.05	0.05
Observations: 80					

Notes: *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. All regressions use heteroskedasticity-robust (HC1) standard errors.

8 Trading Volume Dynamics

In this section, we examine how FOMC communication shocks affect trading activity in futures markets. We focus on abnormal intraday trading volume to capture the immediate response of market participants to newly released policy information. Trading volume reflects both the speed and the magnitude of traders' reactions when new information arrives. Intraday trading volume exhibits pronounced time-of-day patterns. It also varies over time with broader market conditions. As a result, raw volume levels are not directly comparable across intraday intervals or across days. To isolate information-driven trading activity, we construct a measure of abnormal trading volume.

Following the high-frequency event-study literature on monetary policy announcements (Lucca and Moench, 2015), we benchmark realized intraday trading volume against its recent history at the same clock time on non-event days. This normalization removes predictable intraday seasonality and filters out slow-moving variation in trading activity. The resulting measure isolates unexpected trading intensity associated with policy communication. Conceptually, our approach is related to measures of volume shocks defined as deviations from recent averages (e.g., Goyenko et al., 2024). However, our implementation is adapted to a high-frequency intraday setting. This allows the volume response to be precisely aligned with the timing of FOMC statements and press conferences.

Figure 5 provides a descriptive illustration of intraday trading volume around scheduled FOMC announcements. We plot average trading activity on FOMC days relative to non-FOMC days, scaled by a 21-day intraday baseline at the same clock time. The figure shows a sharp increase in trading intensity at the time of the FOMC statement release, followed by elevated activity into subsequent trading sessions, consistent with intensive information processing and heterogeneous interpretation of newly released policy communication.

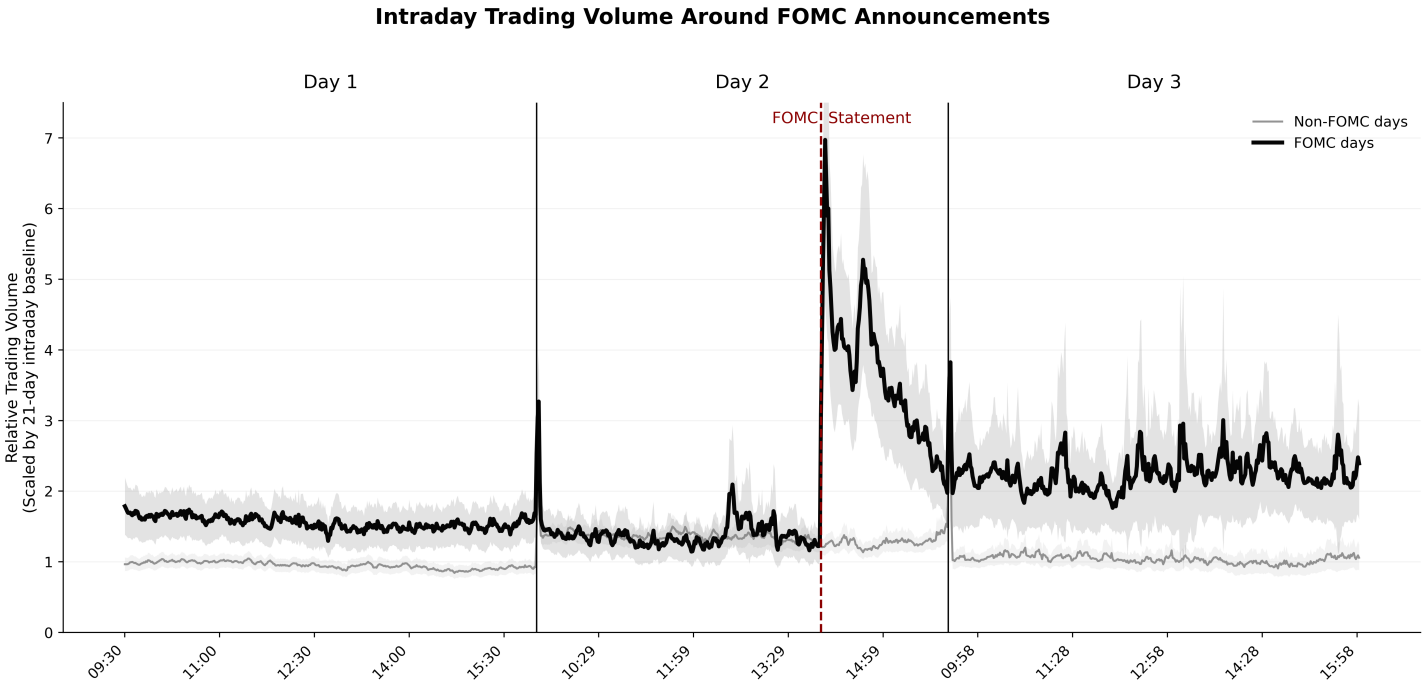


Figure 5: Intraday Trading Volume Around FOMC Announcements. The figure plots average intraday trading volume in E-mini S&P 500 futures on scheduled FOMC days (black) and non-FOMC days (gray), scaled by a 21-day intraday baseline. The dashed vertical line marks the FOMC statement release. Shaded areas denote pointwise 95% confidence intervals.

8.1 Abnormal Intraday Trading Volume

Let $V_{i,t,h}$ denote the trading volume of asset i in intraday interval h on day t. Throughout the analysis, each asset is represented by its front-month futures contract. The index i therefore refers to different underlying asset classes, including federal funds, equity index, and Treasury futures. Intraday time is divided into fixed-length intervals, such as one-minute bins.

Let $\mathcal{H}(w)$ denote a symmetric event window of width $2w$ centered on the announcement time. Total trading volume within this window is defined as

$$V_{i,t}^{(w)} = \sum_{h \in \mathcal{H}(w)} V_{i,t,h}. \tag{19}$$

Intraday trading volume exhibits strong time-of-day patterns and persistent variation

over time. As a result, raw volume levels are not directly comparable across days or intraday intervals. To remove predictable intraday seasonality, we benchmark realized trading volume against activity observed at the same clock time on prior days.

Specifically, we construct a baseline volume using the previous N trading days:

$$\bar{V}_{i,t}^{(w)} = \frac{1}{|\mathcal{D}_t|} \sum_{d \in \mathcal{D}_t} \sum_{h \in \mathcal{H}(w)} V_{i,d,h}, \tag{20}$$

where $\mathcal{D}_t$ denotes the set of recent non-event trading days. We exclude days containing scheduled FOMC announcements and major macroeconomic releases. This restriction ensures that the benchmark reflects normal trading conditions. We set $N = 21$, corresponding to approximately one trading month.

Abnormal trading volume is defined as the log deviation of realized volume from its benchmark:

$$ATV_{i,t}^{(w)} = \log V_{i,t}^{(w)} - \log \bar{V}_{i,t}^{(w)}. \tag{21}$$

This measure captures the percentage increase in trading activity relative to typical conditions at the same time of day. Positive values indicate unusually high trading intensity, while values close to zero indicate normal activity. By construction, this measure removes intraday seasonality and slow-moving variation in trading volume.

8.2 Baseline Regression Specification

We study how post-statement communication shocks affect abnormal intraday trading volume. For each asset i, meeting t, and event window w, we estimate the following baseline specification:

$$ATV_{i,t}^{(w)} = \alpha_i + \beta \, \text{CommShock}_t + \varepsilon_{i,t}^{(w)}. \tag{23}$$

The variable CommShock_t captures the unexpected component of FOMC communication released after the policy decision. Asset fixed effects α_i absorb persistent differences in average trading intensity across futures markets. The coefficient β measures the effect of communication shocks on abnormal trading volume.

A positive estimate of β indicates that unexpected policy communication leads to elevated intraday trading activity. Such a response is consistent with intensified information processing and order flow adjustments following the release of new policy information. The specification focuses on high-frequency trading responses and avoids daily aggregation. This feature allows us to separately identify trading reactions to closely timed communication events, such as FOMC statements and press conferences, that occur on the same calendar day.

8.3 Liquidity and Volatility Controls

A natural concern is that the observed trading responses around FOMC announcements may reflect changes in market liquidity or volatility, rather than information conveyed by FOMC communication itself. Scheduled FOMC meetings often coincide with elevated trading activity, heightened uncertainty, and shifts in risk-bearing capacity. Even in the absence of meaningful informational content, these factors could mechanically generate abnormal volume responses

To address this concern, we examine whether standard liquidity and volatility measures can account for the observed FOMC-day trading patterns. We estimate regressions of abnormal intraday trading volume on an indicator for FOMC announcement days, controlling for both innovations and lagged levels of trading volume and market volatility. Trading volume innovations are constructed as residuals from an AR(1) specification using intraday volume measured at the same clock time on prior non-event days. Volatility innovations are defined analogously using the VIX index. We also consider specifications in which these innovations are interacted with the FOMC-day indicator. This allows liquidity and volatility dynamics to differ systematically on announcement days. These interactions capture the possibility that market conditions respond more strongly to information releases during scheduled policy events.

Table 4 examines whether the abnormal trading activity on FOMC announcement days can be accounted for by standard measures of liquidity and volatility. As a benchmark, column (1) shows that the FOMC-day indicator is economically large and highly statis-

Table 3: Abnormal Trading Volume Response to Communication Shock

	2-Year Treasury Futures	5-Year Treasury Futures	10-Year Treasury Futures	Euro FX Futures	S&P 500 Futures
5-minute window					
CommShock	-7.46^{**}	-8.09^{***}	-4.63^{***}	-7.39^{***}	-4.35^{*}
	(3.60)	(2.84)	(1.73)	(2.11)	(2.22)
R^2	0.03	0.08	0.02	0.04	0.02
10-minute window					
CommShock	-9.82^{***}	-9.88^{***}	-5.36^{***}	-6.85^{***}	-2.63
	(3.14)	(2.61)	(1.70)	(2.17)	(2.21)
R^2	0.04	0.10	0.03	0.03	0.01
30-minute window					
CommShock	-8.20^{***}	-8.97^{***}	-3.80^{**}	-4.67^{**}	-2.15
	(2.80)	(2.89)	(1.48)	(1.95)	(1.93)
R^2	0.03	0.08	0.02	0.02	0.01
60-minute window					
CommShock	-7.80^{***}	-9.11^{***}	-2.91^{*}	-3.59^{*}	-1.60
	(2.43)	(3.30)	(1.48)	(1.93)	(1.67)
R^2	0.03	0.07	0.01	0.01	0.00
90-minute window					
CommShock	-7.12^{***}	-8.34^{***}	-1.86	-2.84	-1.10
	(2.46)	(3.20)	(1.57)	(1.92)	(1.62)
R^2	0.02	0.06	0.01	0.01	0.00
Observations	118	118	118	118	118

Notes: This table reports estimates from regressions of abnormal intraday trading volume on the communication shock. Each column corresponds to a front-month futures contract. Standard errors are reported in parentheses. R^2 denotes the coefficient of determination. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

tically significant. Controlling for trading volume dynamics in column (2) reduces the estimated FOMC-day effect substantially, and innovations in trading volume are strongly significant predictors of abnormal activity. However, the FOMC-day coefficient remains positive and statistically significant, indicating excess trading beyond what is explained by volume controls. Turning to volatility, column (3) shows that VIX innovations also predict abnormal trading, but they have little impact on the estimated FOMC-day effect. When we jointly control for liquidity and volatility in column (4), both trading volume innovations and VIX innovations remain statistically significant. Relative to the benchmark specification in column (1), the coefficient on the FOMC-day indicator declines from about 1.7 to roughly 0.27. This represents a reduction of approximately 80 percent. Even so, the remaining effect is still economically large and highly statistically significant. Finally, column (5) uses innovations constructed from AR(1) specifications augmented with the FOMC-day indicator. The FOMC-day effect remains large and significant, suggesting that conventional liquidity and volatility channels do not explain away the abnormal trading response on FOMC announcement days. Overall, these results imply that liquidity and volatility conditions account for part of the variation in intraday trading activity, but a substantial component of the FOMC-day response remains unexplained by these controls.

8.4 Robustness and Stability Checks

We re-estimate the baseline specification using several out-of-sample approaches. Table 6 reports results from rolling, exponentially weighted, and expanding-window estimations. The goal is to test whether the main effect depends on a specific sample period or estimation window. This exercise evaluates the stability of the baseline findings.

The effect is small and statistically indistinguishable from zero in samples ending in 2015. In contrast, the coefficient becomes strongly negative and statistically significant in samples ending in 2020 and 2025. This pattern holds across all estimation schemes. The magnitude is largest in exponentially weighted regressions. This result suggests that markets place greater weight on recent communication regimes. Overall, the evidence indicates that the impact of communication shocks is not driven by a particular window

Table 4: Liquidity and Volatility Risk Controls

	(1)	(2)	(3)	(4)	(5)
FOMC day dummy	1.696***	0.271***	1.707***	0.265***	1.734***
	[0.071]	[0.042]	[0.072]	[0.042]	[0.039]
Trade Vol (innov.)		0.833***		0.837***	
		[0.008]		[0.008]	
Trade Vol (lag)		0.133***		0.139***	0.145***
		[0.007]		[0.007]	[0.007]
VIX (innov.)			0.036***	0.000	
			[0.008]	[0.003]	
VIX (lag)			0.002	-0.009***	-0.009***
			[0.002]	[0.001]	[0.001]
Trade Vol (FOMC-innov.)					0.837***
					[0.008]
VIX (FOMC-innov.)					0.000
					[0.003]
Const.	0.012	-1.138***	-0.016	-1.032***	-1.125***
	[0.016]	[0.058]	[0.044]	[0.057]	[0.057]
Adjusted R^2	0.09	0.80	0.10	0.80	0.80
Obs.	3,698	3,698	3,698	3,698	3,698

Notes: This table reports estimates from regressions of the dependent variable on an FOMC-day indicator and measures of liquidity and volatility. "Trade Vol (innov.)" is the innovation component of trading volume, and "Trade Vol (lag)" is the lagged level. "VIX (innov.)" and "VIX (lag)" are defined analogously. "Trade Vol (FOMC-innov.)" and "VIX (FOMC-innov.)" are innovations constructed from specifications augmented with the FOMC-day indicator. Robust standard errors are reported in brackets. Significance levels are denoted by ***, **, and * for the 1%, 5%, and 10% levels, respectively.

Table 5: Abnormal Trading Volume: Pooled Panel Regressions

	(1) No FE	(2) Asset FE	(3) Asset + Year FE
5-minute window			
CommShock	-6.142^{***}	-6.142^{***}	-4.427^{***}
	(1.090)	(1.091)	(1.258)
R^2	0.028	0.036	0.247
N	589	589	589
10-minute window			
CommShock	-6.462^{***}	-6.462^{***}	-4.964^{***}
	(1.095)	(1.109)	(1.200)
R^2	0.030	0.037	0.222
N	589	589	589
30-minute window			
CommShock	-4.848^{***}	-4.848^{***}	-3.313^{***}
	(0.935)	(0.943)	(1.027)
R^2	0.019	0.031	0.198
N	589	589	589
60-minute window			
CommShock	-4.207^{***}	-4.207^{***}	-2.861^{***}
	(0.902)	(0.906)	(0.985)
R^2	0.017	0.041	0.214
N	589	589	589
90-minute window			
CommShock	-3.429^{***}	-3.429^{***}	-2.501^{***}
	(0.921)	(0.922)	(0.971)
R^2	0.011	0.044	0.200
N	589	589	589
Asset fixed effects	No	Yes	Yes
Year fixed effects	No	No	Yes

Notes: The dependent variable is abnormal intraday trading volume (ATV) computed for each asset and event window around the scheduled FOMC minutes release. Each block reports pooled OLS estimates for a given post-release window. Heteroskedasticity-robust (HC1) standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

choice and has strengthened in the post-2015 period.

Table 7 examines the time-series relation between communication shocks and abnormal intraday trading volume around FOMC minutes announcements. Column (1) presents the baseline specification. It shows a strong and statistically significant negative association between the communication shock and trading activity. This finding suggests that larger communication surprises are associated with lower intraday liquidity.

Columns (2) through (5) introduce additional controls. These include monetary policy cycle indicators, year fixed effects, macroeconomic variables, and market controls. The estimated coefficient on the communication shock remains negative and statistically significant across all specifications. This result indicates that the main effect is not driven by macroeconomic news, business cycle conditions, or changes in aggregate risk sentiment. Columns (2) and (5) allow for state-dependent effects. They interact the communication shock with tightening and easing indicators. The interaction terms suggest meaningful heterogeneity across policy regimes. Markets appear to respond differently to communication surprises depending on the monetary environment. The results support a robust and economically meaningful link between communication shocks and intraday trading dynamics.

9 Conclusion

This paper examines how financial markets process policy-relevant information that unfolds across sequential Federal Reserve communications. I introduce the Communication Divergence Signal (CDS), a novel measure that combines LLM-derived tone shifts with semantic divergence to capture the joint intensity and novelty of post-meeting communications relative to the initial policy statement. Using a rank-based identification strategy that exploits ordinal information from high-frequency futures surprises, I separate anticipated from unanticipated components of qualitative policy content.

Three main findings emerge. First, communication divergence between FOMC statements and subsequent releases—particularly the minutes—generates economically large and statistically significant asset price responses within narrow intraday windows. Eq-

Table 6: Out-of-Sample Stability of the Communication Shock Effect

	2015	2020	2025
Panel A. 5-year rolling regression			
CommShock ($\hat{\beta}$)	-0.58	-9.20^{**}	-3.75^{*}
	$[-0.54]$	$[-2.07]$	$[-1.78]$
Obs.	1,248	1,248	1,081
Panel B. 10-year rolling regression			
CommShock ($\hat{\beta}$)	-0.25	-3.86^{***}	-5.54^{**}
	$[-0.22]$	$[-3.94]$	$[-2.24]$
Obs.	1,369	2,496	2,329
Panel C. Exponentially weighted regression (half-life = 5 years)			
CommShock ($\hat{\beta}$)	-0.61	-4.87^{***}	-5.76^{***}
	$[-0.60]$	$[-4.38]$	$[-3.36]$
Obs.	1,369	2,617	3,698
Panel D. Exponentially weighted regression (half-life = 10 years)			
CommShock ($\hat{\beta}$)	-0.43	-3.97^{***}	-5.36^{***}
	$[-0.39]$	$[-3.53]$	$[-3.41]$
Obs.	1,369	2,617	3,698
Panel E. Expanding-window regression			
CommShock ($\hat{\beta}$)	-0.25	-3.08^{**}	-4.52^{***}
	$[-0.22]$	$[-2.43]$	$[-3.04]$
Obs.	1,369	2,617	3,698

Notes: This table reports the estimated coefficient on the communication shock, $\hat{\beta}$, across alternative out-of-sample estimation schemes. Columns indicate the *sample ending year*. Panel A (B) estimates the baseline regression using a 5-year (10-year) rolling window ending in the indicated year. Panels C and D apply exponentially decaying weights within the available sample, with half-lives of 5 and 10 years, respectively, placing greater weight on more recent observations. Panel E reports estimates from an expanding-window regression that uses all observations available up to each sample ending year. Adjusted R^2 values range between 0.76 and 0.92 across specifications. Significance levels are denoted by ***, **, and * for the 1%, 5%, and 10% levels, respectively.

Table 7: Abnormal Trading Volume Time Series Regression

	(1) Baseline	(2) + Cycle	(3) + Year FE	(4) + Macro	(5) + Year FE + Macro
CommShock	-17.256^{***}	-33.314^{***}	-40.178^{***}	-30.693^{***}	-36.669^{***}
	(4.779)	(12.737)	(13.119)	(9.885)	(12.746)
tightening		0.163	0.257	0.327	0.457
		(0.215)	(0.354)	(0.211)	(0.427)
easing		0.312	-0.067	0.332	-0.006
		(0.236)	(0.300)	(0.230)	(0.295)
CommShock × tightening		20.930	25.865^{*}	12.428	20.492
		(13.585)	(13.319)	(11.601)	(13.777)
CommShock × easing		20.104^{*}	27.746^{**}	15.897	22.900^{*}
		(12.110)	(12.136)	(9.804)	(12.238)
fff minutes surprise	14.346^{*}	11.891	15.402^{*}	16.914^{**}	16.578^{*}
	(7.499)	(7.576)	(8.791)	(7.107)	(8.517)
VIX innov	0.039	0.033	-0.012	0.021	-0.005
	(0.075)	(0.074)	(0.091)	(0.075)	(0.089)
VIX lag	-0.015	-0.017	-0.021	0.006	-0.008
	(0.017)	(0.017)	(0.028)	(0.020)	(0.030)
$\Delta_{12}\log(\text{CPI})$				-11.816^{**}	-18.592
				(4.891)	(16.043)
$\Delta_{12}\log(\text{IP})$				7.805^{***}	6.729^{*}
				(2.524)	(4.066)
Year fixed effects	No	No	Yes	No	Yes
Cycle dummies	No	Yes	Yes	Yes	Yes
Macro controls	No	No	No	Yes	Yes
Market controls	Yes	Yes	Yes	Yes	Yes
R^2	0.073	0.105	0.242	0.194	0.270
N	115	115	115	115	115

Notes: The dependent variable is abnormal intraday trading volume (ATV) computed over a 10-minute post-release window around the scheduled FOMC minutes announcement. All specifications are estimated by OLS with heteroskedasticity-robust (HC1) standard errors in parentheses. Specifications (2)–(5) allow state dependence via tightening/easing indicators and their interactions with CommShock. Specification (3) includes calendar year fixed effects. Macro controls are year-over-year log changes in CPI and industrial production. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

uity futures decline, Treasury futures rise with effects increasing along the maturity spectrum, and the dollar appreciates following hawkish communication shocks, consistent with markets repricing the expected path of monetary policy. Second, communication shocks are associated with significant changes in abnormal intraday trading volume, indicating that post-meeting releases trigger active information processing and portfolio rebalancing beyond what is explained by liquidity conditions or volatility dynamics. Third, these effects have strengthened markedly in the post-2015 period, suggesting that the informational role of sequential Fed communications has grown as the Fed has expanded its toolkit.

The findings have implications for both policymakers and market participants. For central banks, they underscore that the policy narrative does not end with the statement release. Subsequent communications—minutes, press conferences, and speeches—can shift market expectations in ways that are quantitatively comparable to the initial announcement. This highlights the importance of internal consistency across channels, as unintended divergence may introduce volatility and complicate the transmission of monetary policy. For market participants, the results suggest that systematic monitoring of semantic and tonal shifts across FOMC documents can provide informative signals about the direction and magnitude of market adjustments.

Several limitations should be acknowledged. The LLM-based tone scores are generated using a proprietary model (ChatGPT-4o) whose behavior may vary across versions, raising reproducibility concerns that future work should address by benchmarking against open-source alternatives or dictionary-based methods. The proportional anticipation structure assumed in the identification framework, while parsimonious, does not allow for the possibility that markets may differentially price in tonal versus semantic content. Additionally, the analysis focuses on U.S. markets and Federal Reserve communications; the extent to which these findings generalize to other central banks with different practices remains an open question.

This analysis opens several avenues for further work. First, decomposing the CDS into its tonal and semantic components would allow for a more granular understanding of which dimension drives market responses. Second, extending the framework to incorpo-

rate real-time measures of market attention, such as media coverage or trading activity preceding the release, could refine the separation of priced-in versus novel content. Third, applying the methodology to other central banks, particularly the European Central Bank and the Bank of England, would test whether the informational role of communication sequencing is a general feature of modern monetary policy or specific to the Federal Reserve's institutional design. Finally, as open-source large language models continue to improve, future work can revisit the measurement framework using fully reproducible and versioned models, allowing for exact replication and version control in text-based monetary policy research.

References

Apel, M. and Grimaldi, M. (2012). The information content of central bank minutes. *Riksbank Research Paper Series*, (87).

Bachmann, R., Gödl-Hanisch, I., and Sims, E. R. (2022). Identifying monetary policy shocks using the central bank's information set. *Journal of Monetary Economics*, 128:1–15.

Bauer, M. D. and Swanson, E. T. (2023). A reassessment of monetary policy surprises and high-frequency identification. *Journal of Monetary Economics*, 137:40–59.

Baumgartner, M. and Zahner, J. (2024). Whatever it takes to understand a central banker: Embedding their words using neural networks. *Bank of England Advanced Analytics Conference Paper*. Presented at the Bank of England Advanced Analytics Conference.

BehnamGhader, P., Adlakha, V., Mosbach, M., Bahdanau, D., Chapados, N., and Reddy, S. (2024). Llm2vec: Large language models are secretly powerful text encoders. *arXiv*, arXiv:2404.05961.

Boukus, E. and Rosenberg, J. V. (2006). The Information Content of FOMC Minutes. *Finance and Economics Discussion Series*, (2006-66).

Campbell, J. R., Evans, C. L., Fisher, J. D. M., and Justiniano, A. (2012). Macroeconomic Effects of Federal Reserve Forward Guidance. *Brookings Papers on Economic Activity*, 2012(1):1–80.

Chen, Z., Wang, Y., Wu, Y., and Liu, Z. (2024). Llm2vec: Large language models are secretly powerful text encoders. *arXiv preprint arXiv:2404.05961*.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end? *Handbook of Macroeconomics*, 1(Part A):65–148.

Cieslak, A. and Schrimpf, A. (2019). Non-monetary news in central bank communication. *Journal of International Economics*, 118:293–315.

Curti, F. and Kazinnik, D. (2023). Machine learning for monetary policy analysis: Bert-based central bank communication classification. *Working Paper*. BIS Working Paper No. 1091, https://www.bis.org/publ/work1091.htm.

Doh, T., Song, D., and Yang, S.-K. (2026). Deciphering federal reserve communication via text analysis of alternative fomc statements. *American Economic Journal: Macroeconomics*. Forthcoming.

Doh, T., Song, T., and Yang, S.-K. (2025). Deciphering Federal Reserve communication via text analysis of alternative FOMC statements. *American Economic Journal: Macroeconomics*. Forthcoming.

Ehrmann, M. and Fratzscher, M. (2007). Communication by central bank committee members: Different strategies, same effectiveness? *Journal of the European Economic Association*, 5(2–3):593–606.

Ehrmann, M. and Talmi, J. (2019). Starting from a blank page? Semantic similarity in central bank communication and market volatility. *Review of Financial Studies*, 33(2):448–481.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Gorodnichenko, Y., Pham, T., and Talavera, O. (2023). The voice of monetary policy: Using text analysis to uncover central bank communication. *Journal of Monetary Economics*, 137:40–59.

Goyenko, R., Kelly, B., Moskowitz, T. J., Su, Y., and Zhang, C. (2024). Trading volume alpha. Working Paper NBER Working Paper, National Bureau of Economic Research.

Gurkaynak, R. S., Sack, B., and Swanson, E. T. (2005). Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*, 1(1):55–93.

Gürkaynak, R. S., Sack, B. P., and Swanson, E. T. (2005). Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*, 1(1):55–93.

Handlan, A. (2022). Words as data: Text-based monetary policy surprises. *Working Paper*. Available at SSRN: https://ssrn.com/abstract=4304283.

Hansen, S. and Kazinnik, D. (2023). Generative models and the future of central bank communication. *CEPR Discussion Paper No. 18227*. Available at https://cepr.org/publications/dp18227.

Hansen, S. and McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of Monetary Economics*, 80:116–132.

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

Hubert, P. and Labondance, F. (2018). Central bank tone and the dispersion of views within monetary policy committees. *Journal of International Money and Finance*, 83:33–51.

Jarociński, M. and Karadi, P. (2020). Deconstructing monetary policy surprises: The role of information shocks. *American Economic Review*, 110(9):2760–2797.

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics*, 47(3):523–544.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65.

Lucca, D. O. and Moench, E. (2015). The pre-FOMC announcement drift. *American Economic Review*, 105(8):2527–2558.

Nakamura, E. and Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: The information effect. *Quarterly Journal of Economics*, 133(3):1283–1330.

Reimers, N. and Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using siamese BERT-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Romer, C. D. and Romer, D. H. (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review*, 94(4):1055–1084.

Shah, A., Weber, M., and Ziebart, D. (2023). When central bankers speak in code: Bert embeddings and monetary policy communication. *Working Paper*. Available at SSRN: https://ssrn.com/abstract=4378812.

Swanson, E. T. (2021). Measuring the effects of Federal Reserve forward guidance and asset purchases on financial markets. *Journal of Monetary Economics*, 118:32–53.

Swanson, E. T. and Williams, J. C. (2014). Measuring the effect of the zero lower bound on medium- and longer-term interest rates. *American Economic Review*, 104(10):3154–3185.

Tenreyro, S. and Thwaites, G. (2016). Pushing on a string: US monetary policy is less powerful in recessions. *Review of Economic Studies*, 83(3):1149–1183.

Appendix

A Futures Contract Specifications

Table A1: Futures Contract Specifications by Asset Class

Instrument	Ticker	Contract Size	Expiration Convention
Interest Rate Futures			
Federal Funds Futures	ZQ	\$5,000,000 (avg. monthly rate)	Last business day of contract month
Eurodollar Futures	GE	\$1,000,000	Third Wednesday of contract month
2-Year Treasury Note Futures	ZT	\$200,000	Third Wednesday of contract month
5-Year Treasury Note Futures	ZF	\$100,000	Third Wednesday of contract month
10-Year Treasury Note Futures	ZN	\$100,000	Third Wednesday of contract month
Equity Index Futures			
E-mini S&P 500 Futures	ES	50 $\times$ S&P 500 Index value	Third Friday of contract month
Foreign Exchange Futures			
EUR/USD Futures	6E	€125,000	Third Wednesday of contract month

Notes: This table summarizes the main futures contracts used in the analysis. Contract sizes denote notional values. All expiration conventions follow standard CME quarterly cycles (March, June, September, and December).

The following instruments are used in the analysis:

- *Federal Funds Futures (FFF):* Tick-level transaction data for the `ZQ` contract, sourced from the CME Group. Contracts settle based on the average effective federal funds rate over the contract month. Contracts are quoted as 100 minus the expected monthly average fed funds rate. Each tick record includes a UTC timestamp, trade price, trade volume, and contract identifier.

- *E-mini S&P 500 Futures:* High-frequency trade and quote data for the E-mini S&P 500 futures (ES), the most widely traded equity index futures contract in the U.S. Each contract represents 50 times the value of the S&P 500 index.
- *Eurodollar Futures:* High-frequency trade data for Eurodollar futures (ticker: GE), which settle in cash based on the 3-month U.S. dollar LIBOR on the settlement date, with a notional principal of $1 million.
- *EUR/USD Currency Futures:* High-frequency trade data for euro FX futures (ticker: 6E), traded on the CME. Each contract represents 125,000 euros.
- *Treasury Futures:* High-frequency trade data for the 2-year (ZT), 5-year (ZF), and 10-year (ZN) U.S. Treasury note futures, obtained from CME. These are standardized contracts that represent notional coupon-bearing Treasury securities, settled by physical delivery on a quarterly cycle.

B Intraday Reaction Episodes

The following episodes illustrate how federal funds futures prices respond to different types of FOMC communication events (see Figure 1 in the main text).

(1) *June 15, 2022.* This date corresponds to a regularly scheduled FOMC meeting (panel a). The FOMC announced a 75-basis-point increase in the target range for the federal funds rate, the largest hike since 1994. The front-month futures price dropped immediately, reflecting a sharp upward revision in the implied rate. At 2:30 p.m., Chair Powell began the press conference, emphasizing the Committee's commitment to restoring price stability. Markets absorbed the information quickly, and prices settled soon after.

(2) *July 6, 2022.* This date marks the release of the minutes from the June meeting (panel b). The minutes revealed a broad consensus on maintaining an aggressive stance against inflation. In contrast to the sharp jump at policy announcements, the price reaction was more gradual—futures prices drifted lower over a ten-minute window as market participants processed the additional qualitative context.

(3) *August 26, 2022.* Chair Powell's keynote address at the Jackson Hole Economic Symposium (panel c) caused an immediate drop in the front-month federal funds futures price. Powell stressed the need for "forceful and sustained" action and warned against "premature loosening." Implied futures rates rose by almost 20 basis points within minutes.

C Timestamp Construction

Precise event timing is essential for matching Federal Reserve communications to high-frequency market data. All times are treated as U.S. Eastern Time and converted to UTC using the `America/New_York` timezone.

FOMC statements follow an established release schedule: 2:15 p.m. ET from September 1994 through March 2011; 12:30 p.m. for meetings with press conferences between April 2011 and December 2012; and 2:00 p.m. since January 2013. Minutes are published at 2:00 p.m. ET, three weeks after each meeting. Press conferences begin at 2:15 p.m. (2011–2012) or 2:30 p.m. (2013–present).

Speech release times vary by speaker and venue. Timestamps are extracted from the Federal Reserve's PDF files when available, with the Board's JSON metadata used as a fallback.

D Text Preprocessing Details

In the initial stage of text processing, we removed all stop words (e.g., "a," "is," "the") and deleted non-forecast numbers such as dates and page numbers. We then tokenized the text into smaller units. For sentiment analysis, we used the Loughran and McDonald (2011) sentiment lexicon, a standard dictionary in economics and finance. Contractions were standardized (e.g., "don't" to "do not"), and negation handling ensures that sentences with an odd count of negations receive a reversed sentiment score. We applied lemmatization to simplify words to their root forms (e.g., "went" to "go," "meetings" to "meeting"). Finally, we standardized key economic expressions by converting them into acronyms (e.g., "consumer price index" to "CPI," "gross domestic product" to "GDP").

E Scheduled Release Timestamps

Table E2: Scheduled Release Times of Key Macroeconomic Communications

Document Type	Release Time	Effective Period	Times per Year
FOMC Statement	2:15 p.m.	Sep. 1994 – Mar. 2011	8
	12:30 p.m.	Apr. 2011 – Dec. 2012 (with press conference)	4
	2:00 p.m.	Jan. 2013 – present	8
Press Conference	2:15 p.m.	Apr. 2011 – Dec. 2012	4
	2:30 p.m.	Jan. 2013 – present	8
Meeting Minutes	2:00 p.m.	1997–2004: Thursdays (two noon exceptions: Jul. 2 and Aug. 20, 1998); 2005–present: Tuesdays, three weeks post-meeting	8
Fed Chair Speeches	Unscheduled	2010-present; release time varies by event and location	$\sim$35
PMI (ISM)	10:00 a.m.	First business day of each month	12
Employment Report	8:30 a.m.	First Friday of each month (unless holiday adjusted)	12

Notes: All times are given in Eastern Time (ET). Fed speeches by the Chair and Vice Chair do not follow a fixed release time. For each speech, we extract the release timestamp from the publicly posted transcript or PDF (when available), or infer it from the earliest financial wire coverage. Times are converted to ET.

Table F3: Summary Statistics Around FOMC Statement Days

		Fed Funds Futures		Eurodollar Futures	
Variable	Statistic	FOMC	Non-FOMC	FOMC	Non-FOMC
Volume (contracts)					
	Mean	409.382	190.067	227.458	147.596
	SD	1391.773	767.976	747.511	557.187
	P01	1.0	1.0	1.0	1.0
	P10	3.0	2.0	2.0	2.0
	P50	82.0	30.0	36.0	25.0
	P90	968.300	434.000	518.000	328.000
	P99	5014.000	2428.000	3090.000	1962.000
Absolute price change (bp)					
	Mean	0.415	0.340	6.488	3.838
	SD	3.032	2.495	41.455	31.207
	P01	0.0	0.0	0.0	0.0
	P10	0.0	0.0	0.0	0.0
	P50	0.0	0.0	0.0	0.0
	P90	1.001	0.527	1.570	1.032
	P99	5.750	4.056	137.933	91.325

This table reports summary statistics of 1-minute trading activity around scheduled FOMC statement days. Volume is measured in contracts, and absolute price changes are expressed in basis points. The sample is restricted to regular trading hours (09:00–17:00 Eastern Time) to avoid periods of thin overnight liquidity.

F High-Frequency Futures Data

In Table F3, we present results for both front-month federal funds futures and Eurodollar futures to characterize liquidity and price variability in short-term interest rate markets at high frequency. Several patterns emerge. First, both markets exhibit substantially higher trading volume on FOMC days relative to non-FOMC days, consistent with concentrated information arrival at policy announcements. Second, price variability also increases on announcement days, as shown by larger upper-tail quantiles (e.g., P90 and P99), indicating the presence of large intraday adjustments in expectations.

The two contracts differ in both institutional coverage and sample span. Federal funds futures provide a direct measure of expectations about the target policy rate and are available throughout the full sample used in the main analysis. In contrast, Eurodollar futures extend further back historically but are discontinued after 2023, leading to a slightly shorter and partially overlapping sample. While Eurodollar contracts embed additional term and credit premia relative to federal funds futures, the qualitative patterns in volume and price dispersion are similar across the two markets, supporting the robustness of our high-frequency identification.

G Pre-FOMC Drift and Violatility

This appendix revisits the pre-FOMC drift documented by Lucca and Moench (2015) and subsequent studies. The goal is to validate our event-time construction and benchmark our sample against established empirical patterns in the pre-FOMC drift literature.

Let t denote the scheduled FOMC announcement day. We construct return windows that follow the Lucca–Moench design and are consistent with recent high-frequency event studies. All windows are computed using one-minute S&P 500 futures prices. We consider five event windows.

1. **Overnight:** Open$(t-1)$ to Open(t). This window captures the overnight return leading into the announcement day.

2. **Day:** Open$(t-1)$ to Close(t). This window captures cumulative returns from the

pre-announcement morning through the close of the announcement day.

3. **Gap:** Close$(t-1)$ to Open(t). This window isolates the close-to-open price adjustment immediately before the announcement day.

4. **Post-announcement:** Ann+10m to Close(t). This window captures the intraday response following the scheduled FOMC release.

5. **Two-day:** Open$(t-1)$ to Close$(t+1)$. This window spans two trading days and captures both pre- and post-announcement dynamics.

Following the literature, "Open" and "Close" refer to fixed intraday timestamps chosen to reduce microstructure noise and ensure comparability across days. For scheduled meetings, we define: (i) Open$(t-1)$ as 10 minutes after the Tuesday open, (ii) Open(t) as 5 minutes after the Wednesday open, (iii) Close$(t-1)$ as 5 minutes before the Tuesday close, and (iv) Close(t) as the Wednesday close.[4]

The announcement timestamp Ann corresponds to the scheduled FOMC statement release. We begin the post-announcement window at Ann+10 minutes. This choice allows for immediate price discovery and reduces bid-ask bounce in the first few minutes after the release.

To construct a non-FOMC benchmark while avoiding overlapping return windows, we compute two non-FOMC returns per week. The first spans Monday 10 minutes after the open to Tuesday close. The second spans Wednesday 10 minutes after the open to Thursday close. Weekend returns are excluded. These windows are comparable in duration to the FOMC windows and preserve independence across observations.

For each window, we estimate

$$y_t = \alpha + \beta \cdot \mathbb{1}\{t \in \text{FOMC}\} + \varepsilon_t,$$

where each observation corresponds to a window-level return or volatility measure. The dependent variable y_t denotes the ES futures log return (scaled by 100) or the realized

[4]Results are robust to small perturbations of these offsets (e.g., 5–15 minutes) and to alternative non-overlapping sampling rules.

volatility proxy computed from one-minute returns within the same window. Table G4 reports the estimated coefficients.

Table G4: ES futures returns and realized volatility

	Overnight	Day	Gap	Post-ann.	Two-day
	$(\mathrm{Op}_{t-1} - \mathrm{Op}_t)$	$(\mathrm{Op}_{t-1} - \mathrm{Cl}_t)$	$(\mathrm{Cl}_{t-1} - \mathrm{Op}_t)$	$(\mathrm{Ann{+}10min} - \mathrm{Cl}_t)$	$(\mathrm{Op}_{t-1} - \mathrm{Cl}_{t+1})$
Panel A: ES front-contract returns					
Intercept	0.0590	0.0780	0.0400	0.0160	0.1560
	(2.17)	(2.34)	(2.45)	(1.34)	(3.66)
FOMC dummy	0.1600	0.1940	0.1240	−0.0750	−0.0290
	(1.91)	(1.17)	(2.17)	(-0.77)	(-0.14)
R^2	0.001	0.001	0.002	0.001	0.000
N	1380	1355	1315	1376	1335
Panel B: ES realized volatility					
Intercept	0.9220	1.1620	0.6470	1.0270	1.5120
	(30.91)	(31.62)	(30.18)	(32.64)	(32.37)
FOMC dummy	−0.0860	0.0950	−0.1060	−0.3460	0.1120
	(-1.33)	(1.02)	(-2.55)	(-4.83)	(0.96)
R^2	0.001	0.001	0.003	0.016	0.001
N	1416	1416	1412	1408	1416

Notes: Each column regresses the window return (Panel A) or realized-volatility proxy (Panel B) on an indicator equal to one on scheduled FOMC announcement days. Returns are log returns scaled by 100. Realized volatility is computed from 1-minute ES returns within the same window (e.g., $\sqrt{\sum r_{1m}^2}$). t-statistics are in parentheses and standard errors are Newey–West (HAC).

H Equity Market Responses to FOMC Announcements

All series are aligned in event time such that $t = 0$ corresponds to the announcement release. Let $r_{e,t} = \Delta \log P_{e,t}$ denote the 1-minute log return for event e at minute t, where the event window is defined as $t \in \mathcal{T} = [-60, 120]$ minutes. To remove event-specific drift, returns are demeaned within each event window:

$$\tilde{r}_{e,t} = r_{e,t} - \bar{r}_e, \quad \bar{r}_e = \frac{1}{T_e} \sum_{t \in \mathcal{T}} r_{e,t},$$

where T_e is the number of available observations for event e.

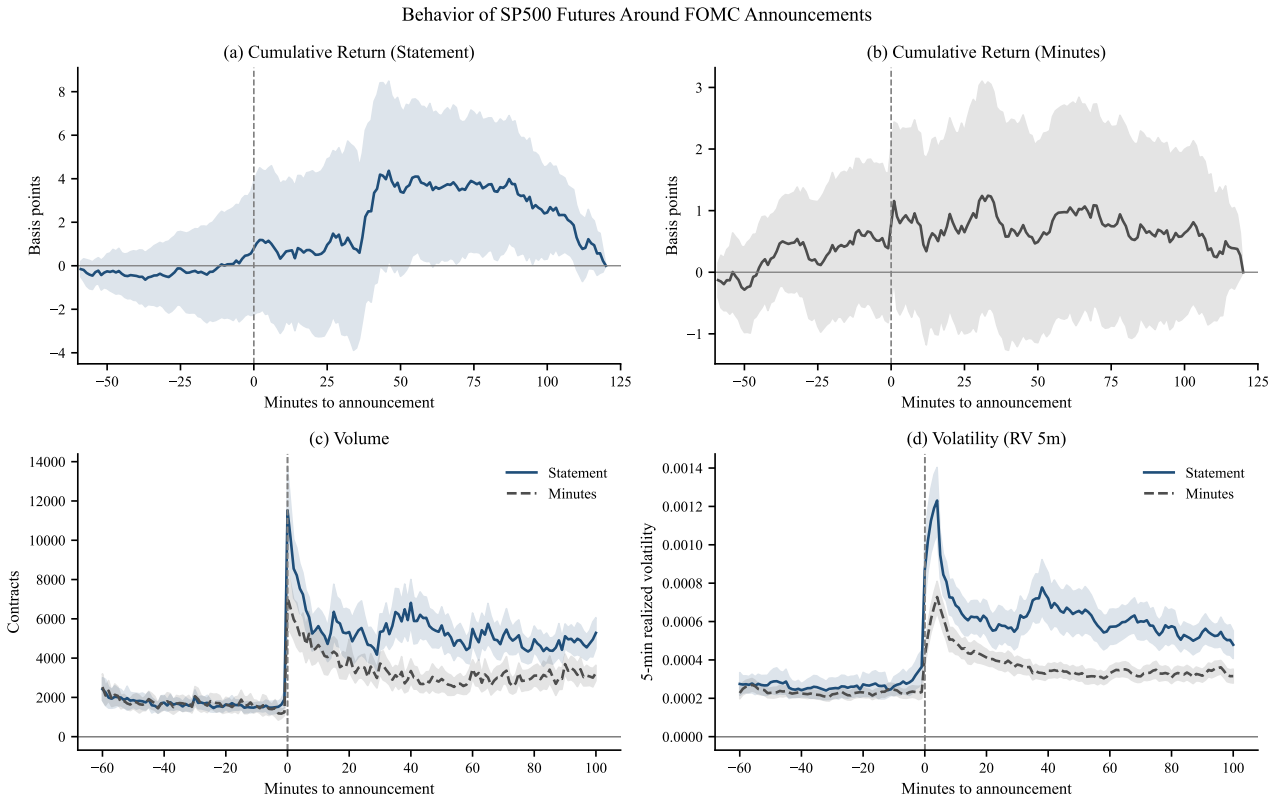


Figure 6: Behavior of E-mini S&P 500 Futures Around FOMC Announcements. Panels (a) and (b) plot the cross-event average cumulative price change of E-mini S&P 500 futures (in basis points) around scheduled FOMC statement releases and FOMC minutes releases, respectively. For each event, the 1-minute price change is demeaned by its event-specific mean (event fixed effect) before cumulation. Panels (c) and (d) report the cross-event average 1-minute trading volume (contracts) and 5-minute realized volatility (standard deviation of log returns), shown separately for statement (solid blue) and minutes (dashed gray) events. Shaded areas denote pointwise 99% confidence intervals computed using the cross-event standard error at each event-time minute. The vertical dashed line marks the announcement time ($t = 0$).

We construct cumulative return paths as

$$C_{e,t} = \sum_{\tau=t_0}^{t} \tilde{r}_{e,\tau}, \quad t_0 = -60,$$

and report the cross-event average

$$\bar{C}_t = \frac{1}{N_t} \sum_{e \in \mathcal{E}_t} C_{e,t},$$

Panels (c) and (d) report cross-event averages of 1-minute trading volume and 5-minute realized volatility. Shaded bands denote pointwise 99% confidence intervals constructed from cross-event standard errors. Specifically, letting s_t denote the cross-event standard deviation,

$$SE_t = \frac{s_t}{\sqrt{N_t}}, \quad \mathrm{CI}_{99\%} = \bar{X}_t \pm 2.576\, SE_t.$$

Realized volatility is computed using a rolling 5-minute window of one-minute log returns. For each event:

$$RV_{e,t}^{(5)} = \sqrt{\frac{1}{5} \sum_{j=0}^{4} \left(r_{e,t-j} - \bar{r}_{e,t}^{(5)} \right)^2}, \quad \bar{r}_{e,t}^{(5)} = \frac{1}{5} \sum_{j=0}^{4} r_{e,t-j}.$$

We then compute the cross-event mean realized volatility:

$$\overline{RV}_t^{(5)} = \frac{1}{N_t} \sum_{e \in \mathcal{E}_t} RV_{e,t}^{(5)}.$$

I Embedding and Semantic Similarity Framework

We use two architectures: (i) a pretrained encoder-based model (SBERT) and (ii) a post-trained decoder adaptation (LLM2Vec). Both approaches represent each FOMC document as a fixed-length embedding vector.

While SBERT specializes in capturing sentence-level paraphrastic similarity, LLM2Vec is specifically designed for long-form policy text. By incorporating bidirectional atten-

tion and contrastive post-training, LLM2Vec effectively captures subtle narrative shifts and contextual dependencies that unfold across the multi-stage FOMC communication process. Because FOMC communications unfold in multiple stages and each stage builds on the context of the previous one, LLM2Vec provides a more effective framework for detecting information updates across communication channels within the same meeting.

I.1 SBERT Model

Algorithm 1 SBERT: Sentence Embedding via Siamese BERT Networks

Inputs: Pretrained BERT parameters θ_0; sentence pairs $\mathcal{D} = \{(x_i, x_i^+)\}_{i=1}^N$; maximum iterations N^{iter}; learning rate α^{learn}; temperature τ.
Output: Fine-tuned encoder parameters θ^* and embedding function $f_{\theta^*}(\cdot)$.

1: Initialize pretrained BERT model with parameters θ_0
2: Create two identical encoder branches $f_\theta^{(1)}$, $f_\theta^{(2)}$ sharing weights θ
3: **for** $i = 1, \ldots, N^{\text{iter}}$ **do**
4: Encode sentence pair (x_i, x_i^+):

$$\mathbf{e}_i = f_\theta^{(1)}(x_i), \quad \mathbf{e}_i^+ = f_\theta^{(2)}(x_i^+)$$

5: Compute cosine similarity:

$$s_i = \frac{\mathbf{e}_i \cdot \mathbf{e}_i^+}{\|\mathbf{e}_i\| \, \|\mathbf{e}_i^+\|}$$

6: Contrastive loss (SimCSE-style):

$$\mathcal{L}_{\text{SBERT}} = -\log \frac{\exp(s_i/\tau)}{\sum_j \exp(s_{ij}/\tau)}$$

7: Update parameters: $\theta \leftarrow \theta - \alpha^{\text{learn}} \nabla_\theta \mathcal{L}_{\text{SBERT}}$
8: **end for**
9: **return** fine-tuned encoder θ^* and embedding function $f_{\theta^*}(\cdot)$

Notes: SBERT employs a Siamese network architecture that encourages high cosine similarity for semantically related pairs and low similarity otherwise.

Model description. Sentence-BERT (SBERT; Reimers and Gurevych, 2019) modifies the BERT architecture to produce fixed-size sentence embeddings suitable for semantic similarity and clustering. Each FOMC sentence is encoded and mean-pooled to obtain a 384-dimensional representation (using the `all-MiniLM-L6-v2` checkpoint from `sentence-transformers`). Embeddings are normalized to unit length and aggregated at the document level via

mean pooling.

I.2 LLM2Vec Model

Algorithm 2 LLM2Vec: Unsupervised Adaptation of Decoder-Only LLMs into Text Encoders

Inputs: Decoder-only LLM parameters θ_0; training corpus $\mathcal{D} = \{x_1, \ldots, x_N\}$; maximum iterations N^{iter}; learning rate α^{learn}; temperature τ.

Output: Fine-tuned encoder parameters θ^* and embedding function $f_{\theta^*}(\cdot)$.

Initialize decoder-only model with θ_0

Replace causal attention mask with bidirectional mask M_{bi}

for $i = 1, \ldots, N^{\text{iter}}$ **do**

  Sample mini-batch $\{x_1, \ldots, x_B\} \subset \mathcal{D}$

  Compute hidden states using M_{bi}

  Randomly mask tokens and predict them via MNTP:

$$\mathcal{L}_{\text{MNTP}} = -\mathbb{E}[\log P(w_{\text{mask}} \mid x_{\setminus \text{mask}}; \theta)]$$

  Encode augmented sentence pairs (x_b, x_b^+) for contrastive learning:

$$\mathcal{L}_{\text{SimCSE}} = -\log \frac{\exp(\text{sim}(\mathbf{e}_b, \mathbf{e}_b^+)/\tau)}{\sum_{b'} \exp(\text{sim}(\mathbf{e}_b, \mathbf{e}_{b'}^-)/\tau)}$$

  Total loss: $\mathcal{L} = \mathcal{L}_{\text{MNTP}} + \lambda \mathcal{L}_{\text{SimCSE}}$

  Update model parameters: $\theta \leftarrow \theta - \alpha^{\text{learn}} \nabla_\theta \mathcal{L}$

end for

return fine-tuned parameters θ^* and encoder $f_{\theta^*}(\cdot)$

Notes: Bidirectional attention enables full-context encoding; MNTP introduces denoising robustness; and SimCSE aligns semantically similar sentences in the embedding space.

Model description. LLM2Vec (BehnamGhader et al., 2024) converts decoder-only LLMs (e.g., Mistral, LLaMA) into text encoders through a lightweight unsupervised post-training procedure. Training proceeds in three steps: (i) enabling bidirectional attention to capture

full context, (ii) masked next-token prediction for denoising, and (iii) contrastive learning (SimCSE) for semantic alignment. We use the `Mistral-7B-LLM2Vec` checkpoint from `sentence-transformers`. All FOMC texts are sentence-tokenized, embedded, and mean-pooled into document-level vectors.

I.3 Semantic Similarity Computation

For both embedding frameworks, each FOMC communication channel is represented by a document-level embedding vector $\mathbf{v}$. Semantic similarity between two channels a and b within the same meeting is computed as:

$$\text{Sim}(\mathbf{v}^{(a)}, \mathbf{v}^{(b)}) = \frac{\mathbf{v}^{(a)} \cdot \mathbf{v}^{(b)}}{\|\mathbf{v}^{(a)}\| \, \|\mathbf{v}^{(b)}\|}, \qquad \text{Div}(v^{(a)}, v^{(b)}) = 1 - \text{Sim}(v^{(a)}, v^{(b)}).$$

The cosine similarity (*Sim*) measures the alignment of two texts in a high-dimensional semantic space, while the divergence measure captures differences in semantic content and informational emphasis across communications. . In this paper, I compute divergence between the FOMC statement and the other post-meeting communication channel (minutes, press conference, and speeches) for the same meeting.