Returns and Volatility Around FOMC Announcements: A High-Frequency Analysis of Policy Tone and Novelty

Simon-Pierre Boucher[1] Marie-Hélène Gagnon[2] Gabriel J. Power[3]

2026-03-06

Abstract

We decompose FOMC statements into policy tone (hawkish/dovish) and informational novelty (departure from previous messaging) and estimate their effects on high-frequency asset returns and volatility. Using 1-minute data for 7 futures contracts across 148 FOMC events (2008–2025), we find that tone predicts directional returns (a one-standard-deviation dovish shift is associated with equity gains that build to about 12 basis points within two hours) while novelty predicts volatility changes (the stance–novelty interaction on VIX persists from 5 to 120 minutes, $t = -5.06$). Policy stance is associated with realized volatility changes in 6 of 7 contracts ($p < 0.01$). Pre-announcement placebo tests and five independent inference methods validate these results. To construct our measures, we train a dual-model ensemble (MiniLM and BERT) on FOMC communications with data-driven PCA-based reference selection.

JEL Classification: E52, E58, G12, G14

Keywords: Monetary policy, FOMC announcements, high-frequency data, textual analysis, market volatility

[1]PhD student in finance, Université Laval, Quebec City QC Canada G1V 0A6, email: simon-pierre.boucher.1@ulaval.ca

[2]Professor of Finance and Research Fellow, CRREP, Université Laval, email: marie-helene.gagnon@fsa.ulaval.ca

[3]IG Wealth Management Chairholder, Professor of Finance and Research Fellow, CRREP and CRIB, Université Laval, email: gabriel.power@fsa.ulaval.ca

1 Introduction

We study how the textual content of Federal Reserve communications affects asset returns and market volatility at high frequency. The Federal Open Market Committee (FOMC)—composed of the seven members of the Board of Governors, the president of the Federal Reserve Bank of New York, and four of the remaining eleven Reserve Bank presidents on a rotating basis—meets eight times per year to set the target federal funds rate. Since 1994, the Committee has released a public statement after each meeting; since 2011, the Chair has also held post-meeting press conferences. Other communications include meeting minutes (released three weeks later) and the Summary of Economic Projections with individual rate forecasts (the "dot plot"). Our analysis focuses exclusively on post-meeting statements, which are the first piece of information available to market participants at a precise, scheduled time. These statements affect U.S. asset prices, global capital flows, and exchange rates (Blinder et al., 2008; Campbell et al., 2012; Wongswan, 2009; Ehrmann et al., 2011).

Our central research question is: *Do different semantic dimensions of FOMC communications—specifically policy tone and informational novelty—affect financial market returns and volatility through distinct economic channels?* If tone captures the directional policy signal (hawkish vs. dovish) while novelty captures how much genuinely new information the statement contains, standard asset pricing theory predicts that these dimensions should have different effects. In the standard framework, asset prices equal expected future cash flows discounted at a rate that reflects policy expectations (Fama, 1970; Bernanke and Kuttner, 2005), so the directional content of the statement—tone—should move expected cash flows and discount rates, and hence returns. Volatility, by contrast, reflects uncertainty and disagreement about how to interpret new information (Veldkamp, 2011; Patton and Verardo, 2012), so the amount of new language—novelty—should affect information processing complexity and interpretive uncertainty, and hence volatility. We test this prediction using 1-minute price data for 7 futures contracts—the E-mini S&P 500 (ES), futures on the Chicago Board Options Exchange Volatility Index (VIX futures, VX), the 10-year (ZN) and 5-year (ZF) Treasury notes, the U.S. Dollar Index (DX), crude oil (CL), and gold (GC)—across 148 FOMC events from 2008 to 2025, estimating panel minute-level regressions, event-level regressions, and local projection impulse response functions (Jordà, 2005). Our identification strategy exploits the exogenous timing of announcements and is validated by pre-announcement placebo tests (Andersen et al., 2003, 2007).

To measure these dimensions, we construct two primary variables: (1) a tone measure capturing the hawkish-dovish spectrum, and (2) a novelty measure quantifying the semantic distance between consecutive FOMC communications. Concretely, we proceed in three steps. First, we convert each statement into a numerical vector (an *embedding*) using two language models of different sizes—MiniLM (which produces 384-dimensional vectors) and BERT (768-dimensional vectors)—and combine their outputs; using two architecturally distinct models guards against findings that are artifacts of a single model. Second, because generic language models are not trained on central bank language, we adapt both models to the corpus of FOMC statements in two stages: an unsupervised stage based on the Transformer-based Sequential Denoising Auto-Encoder (TSDAE) of Wang et al. (2021), which teaches the models Fed-specific vocabulary, followed by a supervised contrastive learning stage on pairs of statements with known relationships. Third, we locate each statement on the hawkish-dovish axis by measuring its proximity to reference statements at the two extremes of the spectrum. Whereas previous studies select these reference statements by hand—introducing researcher degrees of freedom, since results may depend on which dates the researcher chooses—we select them algorithmically using principal component analysis (PCA), making the construction of the semantic axes fully reproducible. Novelty is then measured as the distance between the embeddings of consecutive statements.

FOMC announcements are widely seen as the most important scheduled monetary policy events in global financial markets, as they contribute substantially to shaping expectations about the future path of the U.S. economy. Because markets are forward-looking, surprising announcements are typically followed by sizable adjustments across several markets (Bernanke and Kuttner, 2005; Blinder et al., 2008). Savor and Wilson (2014) show that a disproportionate share of the equity risk premium is earned on macroeconomic announcement days, with FOMC days being particularly important. Brusa et al. (2015, 2019) find that average stock returns and Sharpe ratios on FOMC days are 20–40 times higher than on non-announcement days. Lucca and Moench (2015) document a systematic pre-FOMC drift in equity prices. Nakamura and Steinsson (2018) and Jaroćinski and Karádi (2020) decompose FOMC surprises into policy shocks and information shocks and show that these have different effects on asset prices.

While earlier research focuses exclusively on the measurable policy surprise content of FOMC announcements, typically extracted from federal funds futures (Kuttner, 2001; Bernanke and Kuttner, 2005), a more recent strand of literature has begun to apply textual analysis to extract subtler information from the language itself (Hansen et al., 2018; Shapiro et al., 2022; Schmeling and Wagner, 2019; Gorodnichenko et al., 2023). This is an important step forward, given that the early contributions identify the effect of unexpected rate changes but ignore the qualitative content of the accompanying statement—forward guidance, risk assessments, and descriptions of economic conditions—which Gürkaynak et al. (2005) show can move long-term yields even when the rate decision is fully anticipated. Within this new strand of literature, however, most studies use daily data (Shapiro et al., 2022; Schmeling and Wagner, 2019; Eklund and Kim, 2024), which cannot separate the immediate market reaction to the statement from confounding information that arrives later in the day. Furthermore, the standard approach treats the statement as a one-dimensional object (hawkish vs. dovish), ignoring other dimensions such as how much the statement departs from prior language. In this paper, we address both limitations by using 1-minute data and by decomposing statements into two distinct dimensions: directional tone and informational novelty.

Textual analysis of financial communications has progressed from keyword counting (Bligh and Hess, 2008) and dictionary-based sentiment scoring (Loughran and McDonald, 2011) to transformer-based language models that capture context and word order (Gentzkow et al., 2019; Devlin et al., 2019; Araci, 2019). Within central bank communication, Hayo and Neuenkirch (2010); Apel and Grimaldi (2012) develop systematic measures of FOMC tone and show that these predict monetary policy expectations. Shapiro et al. (2022); Hansen et al. (2018) use natural language processing (NLP) classifiers to place FOMC statements on the hawkish-dovish spectrum, and Eklund and Kim (2024) find that hawkish sentiment measures predict subsequent realizations of the consumer price index (CPI). However, this literature treats FOMC communication as essentially one-dimensional—hawkish versus dovish—and has not examined whether other dimensions of the text, such as how much the language departs from the prior statement, have independent effects on asset prices.

We define informational novelty as the degree to which a new FOMC statement departs from the previous one, measured as the cosine distance between their sentence-transformer embeddings. It is useful to distinguish this concept from three related constructs. *Tone* or *sentiment* refers to the directional content of text (hawkish vs. dovish) and has been studied using dictionaries (Loughran and McDonald, 2011) and classifiers (Shapiro et al., 2022). *Economic policy uncertainty* (EPU), as measured by Baker et al. (2016), is a macro-level index based on newspaper coverage, tax code provisions, and forecaster disagreement. *Monetary policy surprises*, in the tradition of Kuttner (2001), are the unexpected component of the rate decision itself. Our novelty measure differs from all three: it captures changes in *how* information is presented between consecutive statements, independent of both directional stance and the rate decision. A statement can be highly novel yet neutral in tone (e.g., introducing new forward guidance language), or low in novelty yet accompanied by a large rate surprise.

Why should novelty matter for asset prices? Under rational expectations (Muth, 1961; Fama, 1970), only genuinely new information should move prices; repetitive content should already be priced in. But theory also suggests that novel and familiar information may affect returns and volatility differently. Returns respond to revisions in expected cash flows and discount rates—that is, to directional content. Volatility may respond more to disagreement among market participants about the interpretation of new language or to the processing cost of unfamiliar information (Patton and Verardo, 2012; Hu et al., 2013; Veldkamp, 2011). Consistent with this distinction, Manela and Moreira (2017) show that their news-based implied volatility index (NVIX) predicts future market volatility.

Our principal finding is that tone is associated with directional asset returns while novelty is associated with volatility, consistent with two distinct transmission channels. We make two main contributions. First, we develop a decomposition framework for FOMC statements based on a dual-model ensemble (MiniLM and BERT) fine-tuned on the FOMC corpus with data-driven PCA-based reference selection. Second, we provide the first high-frequency evidence on the differential effects of tone versus novelty, using local projection impulse response functions across 7 futures contracts and 148 FOMC events. In addition, we show that these two dimensions operate through distinct channels: tone is associated with fundamental valuations (equity returns building to about 12 basis points per standard deviation within two hours), while novelty is associated with uncertainty resolution (the stance–novelty interaction on VIX futures has a t-statistic of -5.06 and persists from 5 to 120 minutes). Our results are robust to five independent inference methods (Newey–West, wild bootstrap, clustered standard errors, quantile regression, and permutation tests) and are validated through pre-announcement placebo tests.

The remainder of this paper proceeds as follows. Section 2 describes data construction and processing. Section 3 presents our theoretical framework, propositions, and empirical methodology, including our identification strategy and econometric specifications. Section 4 reports our main empirical findings. Section 5 concludes with implications for monetary policy transmission and central bank communication strategy.

2 Data

2.1 FOMC Statement Collection

We collect 451 FOMC communications from the Federal Reserve Board website, spanning 2000–2025. After filtering to policy statements only, our textual corpus contains $N = 217$ FOMC statements, each averaging 3,840 characters. We use the broader 2000–2025 textual sample to train the embedding models and construct the semantic axes, which requires examples spanning multiple policy regimes. For the high-frequency event study, we focus on 148 FOMC events from 2008 to 2025. The 2008 start date reflects two considerations. First, reliable 1-minute futures data became consistently available following the shift to predominantly electronic trading on CME Globex in 2006–2007; pre-2008 high-frequency data for several of our contracts contains gaps and timing errors that would compromise our minute-level identification. Second, the 2008–2025 window encompasses sufficient variation in monetary policy regimes—the zero lower bound (2008–2015), quantitative easing programs, post-crisis normalization (2016–2019), the pandemic response (2020–2021), and the inflation-driven tightening cycle (2022–2025)—across three Fed Chairs (Bernanke, Yellen, Powell) to identify communication effects without sacrificing data quality (Swanson and Williams, 2014; Gürtler and Gürtler, 2010; Blinder et al., 2008). The broader 2000–2025 textual sample used for model fine-tuning starts in 2000 because the FOMC adopted its current practice of issuing post-meeting statements with substantive policy language beginning in 1999–2000; earlier statements were shorter and formulaic, providing too little variation for our NLP models to learn from.

All statements are obtained directly from Federal Reserve official releases archived on the Board of Governors website. We focus exclusively on post-meeting statements rather than meeting minutes, transcripts, or other Fed communications because these statements represent the information that is immediately available to market participants at precise announcement times (Gürkaynak et al., 2005). This temporal precision is necessary for high-frequency identification (Andersen et al., 2003; Rosa, 2013).

We preprocess statements following Gentzkow et al. (2019): we remove headers, footers, voting records, and administrative content, keeping only policy-relevant text. We account for structural breaks in statement format (e.g., the lengthening of statements post-2008) to avoid introducing spurious variation in our semantic measures (Hansen et al., 2018). We cross-reference all statements with the Fed's public archives to verify completeness.

2.2 Dual-Model Embedding Generation

We convert each FOMC statement into a numerical vector using two sentence transformer models, adapted to the FOMC corpus through the training procedure described in Section 3.1.

2.2.1 Model Architectures

1. **MiniLM** (`all-MiniLM-L6-v2`): 33M parameters, 384-dimensional embeddings.

2. **BERT** (`bert-base-uncased` with mean pooling): 110M parameters, 768-dimensional embeddings. A larger model that captures finer contextual distinctions in policy language (Devlin et al., 2019).

We train both models on the FOMC corpus using the two-stage procedure (TSDAE domain adaptation, then MNRL contrastive learning) described in Section 3.1. The training hyperparameters are: batch size 32 (MiniLM) / 16 (BERT), learning rate 1×10^{-5}, 5 (MiniLM) / 4 (BERT) supervised epochs with cosine warmup schedule (10% warmup ratio), and approximately 4,000 training examples per model from six pairing strategies.

2.2.2 Construction of Training Pairs

We construct training pairs using six strategies: (1) overlapping text segments from the same document; (2) consecutive sentence pairs within a document; (3) hawk–dove contrastive pairs classified through keyword banks; (4) temporal proximity pairs from consecutive FOMC meetings; (5) topic-based pairs sharing the same Fed topic (monetary policy, inflation, employment, etc.); and (6) key phrase paraphrases containing identical policy phrases (*"maintain the target range"*, *"decided to* *raise*", etc.). The diversity of pairing strategies prevents the trained models from learning only a single sentiment dimension.

2.2.3 Data-Driven Reference Selection

Previous studies select reference dates manually or construct counterfactual statements by hand. We instead use the hybrid PCA–percentile approach described in Definition 4 to identify semantically extreme statements algorithmically. This procedure eliminates researcher degrees of freedom and produces more stable centroids (5 reference statements per pole versus 2 in typical manual selection). All semantic axes exceed the 0.04 minimum separation threshold. Ensemble separations are: total policy stance (0.302), risk assessment (0.274), policy communication (0.248), and economic priority (0.252).

We validate the embeddings in three ways. First, we compare embedding-based tone rankings with expert classifications and verify that known dovish and hawkish statements sort correctly. Second, we measure inter-model agreement: mean ensemble confidence is 0.837 for novelty and 0.562–0.806 for tone across axes. Third, we verify that the PCA axis assignments produce intuitive orderings.

2.3 High-Frequency Financial Data Construction

We use 1-minute OHLCV data for seven futures contracts spanning the major asset classes relevant to monetary policy transmission:

- **ES**: E-mini S&P 500 (equity benchmark)—responds to policy communications through discount rate effects, growth expectations, and risk premium adjustments.
- **VX**: VIX Futures (volatility)—captures implied volatility and uncertainty resolution, providing the most direct measure of information processing effects.
- **ZN**: 10-Year Treasury Note—reflects the interaction between policy expectations and term premium effects.
- **ZF**: 5-Year Treasury Note—primarily captures expectations about near-term policy rate changes.

- **DX**: Dollar Index—reflects relative monetary policy stances and international spillover effects.
- **CL**: Crude Oil WTI—reflects both inflation expectations and growth concerns influenced by monetary policy.
- **GC**: Gold—a safe-haven asset whose demand falls when accommodative communication triggers a rotation toward risk assets, and rises with policy-induced uncertainty.

The data span 148 FOMC events (approximately 8 per year from 2008 to 2025). For each event, we extract a ± 120 minute window around the 14:00 ET announcement time. We interpolate the raw data to a regular 1-minute grid with forward-filling of short gaps (≤ 5 minutes) and returns recomputed on the regularized grid. We require at least 80% valid observations per rolling window.

For each instrument, we construct minute-by-minute log returns as:

$$r_{i,m}^{(j)} = \log\left(P_{i,m}^{(j)}\right) - \log\left(P_{i,m-1}^{(j)}\right) \tag{1}$$

where $P_{i,m}^{(j)}$ is the price of asset i at minute m relative to FOMC event j. We align timestamps across instruments and exchanges, handle market closures and trading halts, and filter outliers following standard microstructure procedures.

2.4 Event Window Specification and Temporal Alignment

Our primary analysis employs a 45-minute event window following each FOMC announcement, spanning from the announcement time ($t = 0$) to 45 minutes after ($t = +45$). We choose this window length for two reasons.

First, most of the announcement effect occurs within 15 minutes of the release, but equity responses to tone continue to develop through 45 minutes post-announcement. Second, press conferences typically begin 30 minutes after the statement, so a 45-minute window includes only the first 15 minutes of the press conference, limiting contamination.

We also implement several alternative window specifications as robustness checks. A narrow 15-minute post-announcement window provides very clean identification but with reduced statistical power. An extended 60-minute post-announcement window captures longer adjustment dynamics but with increased contamination risk, particularly from press conferences that typically begin 30 minutes after the statement release.

FOMC statements are typically released at 2:00 PM Eastern Time. We verify actual release times using Federal Reserve timestamps, news services, and market data providers, and correct any timing discrepancies before aligning with price data.

3 Methodology

This section describes our NLP framework for measuring tone and novelty (Section 3.1), presents the econometric specifications we use to estimate the effects of these measures on asset returns and volatility (Section 3.2), and states the testable hypotheses together with the specific coefficient restrictions that each one implies (Section 3.3).

3.1 Theoretical Framework

3.1.1 Conceptual Foundation

We decompose each FOMC statement into two dimensions: policy tone (hawkish vs. dovish) and informational novelty (distance from the previous statement). To illustrate why context-aware methods are needed for this decomposition, consider the difference between dictionary-based and machine learning approaches. A dictionary approach aggregates sentiment as:

$$S_{\text{dict}} = \frac{1}{N} \sum_{i=1}^{N} \nVdash_{[\text{word}_i \in D_{\text{pos}}]} - \nVdash_{[\text{word}_i \in D_{\text{neg}}]} \tag{2}$$

where D_{pos} and D_{neg} are predetermined word lists. The dictionary approach counts words one at a time, without regard to their position or neighbors. A transformer-based model instead produces $S_{\text{ML}} = f_\theta(x_1, x_2, \ldots, x_N; C)$, where f_θ is a neural network with learned parameters θ, and x_i are token representations that depend on surrounding words and document-level context C. To see why context matters, note that "the policy remains accommodative" and "the policy is no longer accommodative" contain the same keyword ("accommodative") but have opposite meanings; a dictionary assigns similar scores to both, while a transformer assigns opposite scores.

Our approach builds on three literatures. The information processing literature predicts that markets react more strongly to genuinely new information than to repetition (Grossman and Stiglitz, 1980; Fama, 1970). The central bank communication literature shows that policy statements move asset prices independently of rate decisions (Gürkaynak et al., 2005; Campbell et al., 2012). The market microstructure literature demonstrates that high-frequency data can isolate the real-time incorporation of new information into prices (Andersen et al., 2003; Rosa, 2013).

Definition 1 (FOMC Statement Representation)**.** *Each FOMC statement F_t released at time t is mapped to a vector $\mathbf{f}_t \in \mathbb{R}^d$ in d-dimensional semantic space through domain-specific transformer embeddings. We employ two architecturally distinct models: MiniLM ($d = 384$) and BERT ($d = 768$), combined into a separation-weighted ensemble.*

Each model maps a statement to a real-valued vector in which semantically similar texts (e.g., "accommodative" and "easy policy") are represented by nearby points, enabling cosine similarity comparisons and PCA. Unlike bag-of-words representations, where each coordinate is a word count and most are zero, these vectors are dense: every coordinate carries information. The output dimensionality ($d = 384$ for MiniLM, $d = 768$ for BERT) is fixed by the model architecture—MiniLM-L6-v2 has 6 layers and 384 hidden units per layer; BERT-base has 12 layers and 768 hidden units (Devlin et al., 2019)—and is not a hyperparameter we select. We adapt both architectures to the FOMC domain using a two-stage training procedure.

Stage 1: TSDAE Unsupervised Pre-Training. We first apply Transformer-based Sequential Denoising Auto-Encoder (TSDAE) (Wang et al., 2021) to adapt each model to the FOMC domain vocabulary:

$$\mathcal{L}_{\text{TSDAE}} = -\sum_{i=1}^{N} \log P\big(x_i \mid \tilde{x}_i; \theta_{\text{enc}}, \theta_{\text{dec}}\big), \tag{3}$$

where $\tilde{x}_i$ is a corrupted version of sentence x_i (token deletion with probability 0.6). This unsupervised stage teaches domain-specific vocabulary and structure without requiring labeled pairs, training for 3 epochs (MiniLM) or 2 epochs (BERT) with learning rate 3×10^{-5}.

Stage 2: Supervised Contrastive Learning with MNRL. We then fine-tune using Multiple Negatives Ranking Loss (MNRL):

$$\mathcal{L}_{\text{MNRL}} = -\frac{1}{B}\sum_{i=1}^{B} \log \frac{\exp\big(\text{sim}(a_i, p_i)/\tau\big)}{\sum_{j=1}^{B} \exp\big(\text{sim}(a_i, p_j)/\tau\big)}, \tag{4}$$

where (a_i, p_i) are positive pairs, $\text{sim}(\cdot, \cdot)$ denotes cosine similarity, τ is the temperature parameter, and B is the batch size. In-batch negatives provide contrastive signal without explicit negative sampling. We construct approximately 4,000 training pairs per model using the six pairing strategies described in Section 2.

Dual-Model Ensemble. The two models are combined using separation-weighted averaging:

$$\widehat{\text{Tone}}^{\text{ens}}_{t,k} = w_k^{\text{MiniLM}} \cdot \text{Tone}^{\text{MiniLM}}_{t,k} + w_k^{\text{BERT}} \cdot \text{Tone}^{\text{BERT}}_{t,k}, \tag{5}$$

where the weights are proportional to axis separation quality:

$$w_k^m = \frac{\text{Sep}_k^m}{\text{Sep}_k^{\text{MiniLM}} + \text{Sep}_k^{\text{BERT}}}, \quad m \in \{\text{MiniLM}, \text{BERT}\}. \tag{6}$$

For each statement, we measure inter-model confidence as:

$$\text{Conf}_{t,k} = \frac{1}{2}\big[\mathbb{1}[\text{sign}(\text{Tone}^{\text{M}}_{t,k}) = \text{sign}(\text{Tone}^{\text{B}}_{t,k})] + (1 - |\text{Tone}^{\text{M}}_{t,k} - \text{Tone}^{\text{B}}_{t,k}| / \max_j |\text{Tone}^{\text{M}}_{j,k} - \text{Tone}^{\text{B}}_{j,k}|)\big]. \tag{7}$$

Mean ensemble confidence is 0.837 for novelty and ranges from 0.562 to 0.806 for tone across axes. This dual-model approach reduces model-specific biases: MiniLM (33M parameters, 384 dimensions) provides computational efficiency and strong separation on policy stance, while BERT (110M parameters, 768 dimensions) captures richer contextual relationships. The ensemble achieves balanced performance across all semantic dimensions.

3.1.2 Semantic Dimensions

Definition 2 (Semantic Similarity Measure)**.** *The semantic similarity between text vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *is quantified using cosine similarity:*

$$sim(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}||_2 \cdot ||\mathbf{v}||_2} = \frac{\sum_{i=1}^{d} u_i v_i}{\sqrt{\sum_{i=1}^{d} u_i^2} \cdot \sqrt{\sum_{i=1}^{d} v_i^2}} \tag{8}$$

The cosine similarity measure ranges from -1 to 1, where values approaching 1 indicate high semantic similarity, values near 0 indicate orthogonal or unrelated content, and values approaching -1 indicate semantic opposition. This metric is particularly well-suited for high-dimensional text embeddings because it captures angular distance between vectors while being invariant to vector magnitude differences that might arise from statement length variations.

Definition 3 (Statement Novelty)**.** *The informational novelty of statement* F_t *relative to the immediately preceding statement is defined as:*

$$Novelty_t = 1 - sim(\mathbf{f}_t, \mathbf{f}_{t-1}) \tag{9}$$

This novelty measure captures the degree to which current policy communications represent genuine departures from recent messaging patterns. Higher novelty values indicate statements containing substantively new informational content relative to recent communications, while lower values suggest continuity with established messaging. The measure is bounded between 0 and 2, where 0 indicates identical statements and 2 indicates maximally opposed statements.

To construct our policy tone measure, we develop a data-driven approach based on PCA-based semantic axis construction that replaces manually selected reference dates with an algorithmic procedure, eliminating researcher degrees of freedom.

Definition 4 (Data-Driven Semantic Axis Construction)**.** *Let* $\mathbf{E} \in \mathbb{R}^{N \times d}$ *be the L2-normalized embedding matrix. We compute the first* $K = 8$ *principal components:*

$$\mathbf{E} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top + \boldsymbol{\epsilon}. \tag{10}$$

For each semantic dimension $k \in \{1, \ldots, 4\}$ *(policy stance, risk assessment, policy communication,* *economic priority), we define keyword banks $\mathcal{K}_k^+$ (positive pole) and $\mathcal{K}_k^-$ (negative pole), and compute a keyword differential score:*

$$\Delta_i^{(k)} = \sum_{w \in \mathcal{K}_k^+} \mathbb{1}[w \in text_i] - \sum_{w \in \mathcal{K}_k^-} \mathbb{1}[w \in text_i]. \tag{11}$$

We match each semantic axis to the PC maximizing the absolute Pearson correlation:

$$j^*(k) = \arg\max_{j \in \{1,\dots,K\}} \left| Corr(\mathbf{v}_j, \boldsymbol{\Delta}^{(k)}) \right|. \tag{12}$$

The reference selection uses a hybrid seed-and-percentile approach: (1) use 3 expert-selected dates per pole to define an initial axis direction $\hat{\mathbf{d}}_k = \bar{\mathbf{e}}_{\text{pos}} - \bar{\mathbf{e}}_{\text{neg}}$; (2) project all embeddings onto this direction: $s_i = \mathbf{e}_i^\top \hat{\mathbf{d}}_k / \|\hat{\mathbf{d}}_k\|$; (3) select $M = 5$ statements with highest and lowest projection scores; (4) iteratively refine by recomputing the axis direction from selected references and re-selecting. This yields 5 reference statements per pole (compared to 2 in manual selection), producing more stable centroids.

For each axis k, we measure discrimination quality using the separation metric:

$$\text{Sep}_k = 1 - \cos(\bar{\mathbf{c}}_k^+, \bar{\mathbf{c}}_k^-), \tag{13}$$

where $\bar{\mathbf{c}}_k^+, \bar{\mathbf{c}}_k^-$ are the L2-normalized positive and negative centroids. All axes achieve separations well above the 0.04 minimum threshold, with ensemble separations ranging from 0.248 to 0.302.

Definition 5 (Policy Tone)**.** *The directional tone of statement F_t on semantic axis k is measured as:*

$$Tone_{t,k} = \frac{\cos(\mathbf{f}_t, \bar{\mathbf{c}}_k^+) - \cos(\mathbf{f}_t, \bar{\mathbf{c}}_k^-)}{Sep_k}, \quad \textit{clipped to } [-1, +1]. \tag{14}$$

This tone measure captures the relative proximity of actual statements to the positive (hawkish) and negative (dovish) reference centroids for each semantic axis, normalized by the axis separation. The primary variable used in regressions is the *total policy stance tone*. In the regression tables, the z-scored stance variable is oriented so that positive values indicate more accommodative (dovish) communication; positive coefficients therefore measure the response to a one-standard-deviation dovish shift. The composite MPS measure defined below retains the hawkish-positive orientation.

Relative to counterfactual-based methods, this approach is reproducible, does not depend on particular date choices, and extends to semantic dimensions other than the hawkish-dovish spectrum.

3.1.3 Integrated Policy Stance Measurement

Definition 6 (Monetary Policy Stance)**.** *The overall monetary policy stance is defined as:*

$$Stance_t = Novelty_t \times Tone_t \tag{15}$$

Note on originality: Definitions 1–11 and Proposition 1 are our own contributions. They apply the standard sentence-embedding framework from the NLP literature to monetary policy communications with a new decomposition structure.

The multiplicative form means that stance is large only when a statement is both novel *and* directionally clear. We adopt a multiplicative rather than additive form for two reasons. First, it ensures that Stance is zero whenever either component is zero: repetitive content has no impact regardless of its tone, and novel but directionally neutral content has none either. Second, it captures the reinforcing interaction between the two dimensions: novel information amplifies the impact of tone, and clear directional tone amplifies the impact of novelty.

In the following two definitions, $\mathbf{f}_t^D$ and $\mathbf{f}_t^H$ denote the embeddings of counterfactual statements located at the dovish and hawkish reference centroids of the policy stance axis, $\bar{\mathbf{c}}^-$ and $\bar{\mathbf{c}}^+$; they represent what the Fed would have released had it issued a maximally dovish or maximally hawkish statement at meeting t.

Definition 7 (Pure Dovish and Hawkish Stances)**.** *The extreme stance scenarios are defined as:*

$$Stance_t^{dove} = -\left(1 - sim(\mathbf{f}_t^D, \mathbf{f}_{t-1})\right) \tag{16}$$

$$Stance_t^{hawk} = 1 - sim(\mathbf{f}_t^H, \mathbf{f}_{t-1}) \tag{17}$$

Definition 8 (Weighted Policy Stance Representation)**.** *The overall stance can be expressed as a convex combination of extreme stances:*

$$Stance_t = w_t \cdot Stance_t^{dove} + (1 - w_t) \cdot Stance_t^{hawk} \tag{18}$$

where the dovish weight parameter w_t is computed (not estimated) as a deterministic function of the observed similarities:

$$w_t = \frac{1 - sim(\mathbf{f}_t^H, \mathbf{f}_{t-1}) - (1 - sim(\mathbf{f}_t, \mathbf{f}_{t-1})) \cdot Tone_t}{2 - sim(\mathbf{f}_t^D, \mathbf{f}_{t-1}) - sim(\mathbf{f}_t^H, \mathbf{f}_{t-1})} \tag{19}$$

This weighted representation provides an alternative interpretation of policy stance as a position along the spectrum between extreme dovish and hawkish alternatives. The weight parameter w_t is not a free parameter to be estimated; rather, it is algebraically derived from the observed textual similarities.

3.1.4 Market Surprise Identification

Definition 9 (Novelty Decomposition)**.** *The total novelty can be decomposed into predictable and unpredictable components:*

$$Novelty_t = \overline{Novelty}_{t|t-\Delta} + \varepsilon_t \tag{20}$$

where $\overline{Novelty}_{t|t-\Delta} \equiv \mathbb{E}_{t-\Delta}[Novelty_t]$ represents the expected novelty based on information available Δ periods before the announcement, and ε_t represents the novelty surprise component with $\mathbb{E}_{t-\Delta}[\varepsilon_t] = 0$ and $Var(\varepsilon_t) = \sigma_\varepsilon^2 < \infty$. We assume ε_t is covariance-stationary with $Cov(\varepsilon_t, \varepsilon_{t-j}) \to 0$ as $j \to \infty$.

Definition 10 (Expected Policy Stance)**.** *Under rational expectations, the expected policy stance is:*

$$\mathbb{E}_{t-\Delta}[Stance_t] = \mathbb{E}_{t-\Delta}\left[Novelty_t \times Tone_t\right] \tag{21}$$

Assumption 1 (Conditional Independence or Zero Covariance)**.** *We assume either (i) conditional independence: $Novelty_t \perp Tone_t | \mathcal{I}_{t-\Delta}$, where $\mathcal{I}_{t-\Delta}$ is the information set at time $t-\Delta$, or (ii) the weaker condition of zero conditional covariance:*

$$Cov_{t-\Delta}(Novelty_t, Tone_t) = 0 \tag{22}$$

This assumption is plausible because novelty captures whether the Fed changes its messaging structure (how information is presented), while tone captures the directional policy stance (hawkish/dovish). Empirically, the correlation between novelty and tone in our sample is 0.12 (Figure 2), supporting the assumption that these are largely independent dimensions. Under this assumption:

$$\mathbb{E}_{t-\Delta}[\text{Stance}_t] = \mathbb{E}_{t-\Delta}[\text{Novelty}_t] \times \mathbb{E}_{t-\Delta}[\text{Tone}_t] \tag{23}$$

Further assuming, for tractability, that tone is binary ($\text{Tone}_t \in \{-1, +1\}$, taking the dovish value -1 with probability $p_{t-\Delta}$), so that $\mathbb{E}_{t-\Delta}[\text{Tone}_t] = 1 - 2p_{t-\Delta}$:

$$\mathbb{E}_{t-\Delta}[\text{Stance}_t] = (1 - 2p_{t-\Delta}) \cdot \overline{\text{Novelty}}_{t|t-\Delta} \tag{24}$$

Definition 11 (Policy Stance Surprise)**.** *The communication-based policy stance surprise is defined as:*

$$MPS_t = Stance_t - \mathbb{E}_{t-\Delta}[Stance_t] \tag{25}$$

We use the abbreviation MPS (monetary policy surprise) for brevity, but emphasize that this measure is conceptually distinct from the traditional rate-based monetary policy surprise of Kuttner (2001), which measures the unexpected component of the federal funds rate decision using futures prices. Our MPS is a *communication* surprise—it captures unexpected variation in the textual content of FOMC statements, not in the quantitative rate decision. The two measures are complementary and potentially orthogonal: a meeting can produce a large rate surprise with a routine statement, or a novel and directional statement with a fully anticipated rate decision. Our measure captures the latter dimension, which has become increasingly important as forward guidance and qualitative communication have gained prominence in the monetary policy toolkit.

Proposition 1 (MPS Decomposition Structure)**.** *The monetary policy surprise admits the following decomposition:*

$$MPS_t = \overline{Novelty}_{t|t-\Delta}(Tone_t + 2p_{t-\Delta} - 1) + Tone_t \cdot \varepsilon_t \tag{26}$$

This decomposition reveals that policy surprises consist of two distinct components: (1) surprises arising from unexpected tone conditional on expected novelty, and (2) surprises arising from unexpected novelty, weighted by the actual tone of the communication. The proof of Proposition 1, together with the derivations underlying Definitions 7 and 8, is in Appendix A.

3.2 Econometric Framework

3.2.1 Panel Data Structure

The data are organized as a three-dimensional panel structure with observations indexed along three dimensions:

- Asset identifier $i \in \{1, 2, \ldots, I\}$ where I represents the total number of financial instruments
- Event identifier $j \in \{1, 2, \ldots, J\}$ corresponding to FOMC announcement dates where $J = 148$
- Relative minute $m \in \{-120, \ldots, +120\}$ measured from announcement time, with the primary post-announcement window covering $m \in \{0, \ldots, +45\}$

3.2.2 Identification Strategy

Our identification relies on the fact that FOMC statements are drafted and finalized before the announcement and released at a scheduled time (typically 2:00 PM ET). The text at $t = 0$ therefore cannot be influenced by market reactions at $t > 0$, satisfying the timing restriction for causal interpretation.

Our identification relies on three core conditions:

Identification Condition 1 (Exogenous Timing). The precise minute-by-minute timing of FOMC announcements is predetermined and exogenous to short-term market movements within our event windows.

Identification Condition 2 (Information Concentration). Within our 45-minute post-announcement event windows, FOMC statements represent the dominant source of monetary policy-relevant information, with minimal contamination from other systematic news sources.

Identification Condition 3 (Market Efficiency). Financial markets rapidly incorporate new information from FOMC statements into prices, enabling causal interpretation of immediate price movements following announcements.

We acknowledge two potential threats to this identification strategy. First, the Federal Reserve may adjust its communication in response to broader financial conditions observed before the meeting. We mitigate this concern by including event fixed effects δ_j that absorb all meeting-level heterogeneity. Second, FOMC announcements may coincide with other information releases (e.g., Summary of Economic Projections, press conference expectations). Our focus on the immediate post-announcement window limits contamination from subsequent press conference content, which typically begins 30 minutes after the statement release.

3.2.3 Dynamic Response Function Estimation

Our primary empirical specification estimates dynamic impulse response functions:

$$r_{i,m}^{(j)} = \alpha_i + \delta_j + \sum_{k=0}^{+45} \beta_k \cdot \mathbb{1}_{[m=k]} \cdot \mathrm{MPS}_j + \varepsilon_{i,m,j} \tag{27}$$

where $r_{i,m}^{(j)}$ denotes the log return of asset i at minute m relative to monetary policy announcement j, α_i represents asset fixed effects, δ_j captures event fixed effects, and β_k measures the marginal impact of monetary policy surprises occurring exactly k minutes relative to the announcement.

3.2.4 Semantic Decomposition Specification

To examine the differential effects of policy tone versus informational novelty, we estimate:

$$r_{i,m}^{(j)} = \alpha_i + \delta_j + \sum_{k=0}^{+45} \left[\beta_k^{(T)} \cdot \mathbb{1}_{[m=k]} \cdot \mathrm{Tone}_j + \beta_k^{(N)} \cdot \mathbb{1}_{[m=k]} \cdot \mathrm{Novelty}_j\right] + \varepsilon_{i,m,j} \tag{28}$$

The tone coefficients $\{\beta_k^{(T)}\}$ measure the dynamic response to hawkish versus dovish content, while the novelty coefficients $\{\beta_k^{(N)}\}$ capture the response to informational content, independent of directional bias.

3.2.5 Rolling Realized Measures

In addition to return effects, we examine how policy communications affect market volatility using rolling realized measures computed from 1-minute log returns.

Realized Variance. We construct NA-tolerant rolling realized variance in K-minute windows:

$$\mathrm{RV}_K(t) = \frac{K}{n_{\mathrm{valid}}} \sum_{s=t-K+1}^{t} r_s^2 \cdot \mathbb{1}[r_s \text{ valid}], \tag{29}$$

where n_{valid} is the number of non-missing observations in the window. We require at least 80% valid observations.

Realized Beta. We compute rolling beta relative to the ES (S&P 500) benchmark:

$$\hat{\beta}_K(t) = \frac{\sum_{s=t-K+1}^{t} r_{i,s} \cdot r_{m,s}}{\sum_{s=t-K+1}^{t} r_{m,s}^2}, \tag{30}$$

where $r_{m,s}$ denotes the ES benchmark return. We use $K = 5$ for panel minute-level regressions and $K = 30$ for event-level analysis.

3.2.6 Panel Minute-Level Regressions

For each ticker, we estimate:

$$\log(\text{RV}_{i,t}) = \alpha + \beta \cdot (\text{Post}_t \times \text{Semantic}_i) + \varepsilon_{i,t}, \tag{31}$$

where $\text{Post}_t = \mathbb{1}[t > 0]$ indicates post-announcement minutes, and Semantic_i is either *Stance* or *Novelty* (z-scored). Standard errors are clustered by event date.

3.2.7 Event-Level Regressions

For each event, we compute pre/post changes:

$$\Delta \log(\text{RV})_i = \overline{\log(\text{RV})}_{\text{post}} - \overline{\log(\text{RV})}_{\text{pre}}, \tag{32}$$

and regress on semantic measures:

$$\Delta Y_i = \alpha + \beta_1 \text{Stance}_i + \beta_2 \text{Novelty}_i + \beta_3 (\text{Stance}_i \times \text{Novelty}_i) + \varepsilon_i, \tag{33}$$

with Newey–West HAC standard errors. We estimate this specification for five dependent variables: ΔRV, $\Delta \log(\text{RV})$, RV ratio, log(RV ratio), and $\Delta\beta$.

3.2.8 Local Projection Impulse Response Functions

Following Jordà (2005), for each horizon $h \in \{1, 2, 3, 5, 10, 15, 20, 30, 45, 60, 90, 120\}$ minutes post-announcement:

$$\text{CumRet}_i(0 \to h) = \alpha + \beta_1^{(h)}\text{Stance}_i + \beta_2^{(h)}\text{Novelty}_i + \beta_3^{(h)}(\text{Stance}_i \times \text{Novelty}_i) + \varepsilon_i^{(h)}. \tag{34}$$

Abnormal cumulative returns are computed as $\text{CAR}_i(h) = \text{CumRet}_i(h) - \text{CumRet}_i^{\text{ES}}(h)$. This local projection approach is robust to misspecification of the data-generating process and allows horizon-specific inference without imposing a parametric impulse response shape. Because the post-meeting press conference typically begins 30 minutes after the statement release, we interpret responses at horizons beyond $h = 30$ as the joint effect of the statement and early press-conference communication.

3.2.9 Pre-Announcement Placebo Test

We estimate the same specification (34) for h minutes *before* the announcement:

$$\text{CumRet}_i(-h \to 0) = \alpha + \beta_1^{(-h)}\text{Stance}_i + \beta_2^{(-h)}\text{Novelty}_i + \beta_3^{(-h)}(\text{Stance}_i \times \text{Novelty}_i) + \varepsilon_i^{(-h)}. \tag{35}$$

Under the null that announcement content is not anticipated, all $\beta^{(-h)}$ should be zero. This placebo test provides a direct validation of our event study design.

3.2.10 Statistical Inference and Robustness

We subject all event-level results to five alternative inference methods:

1. **Newey–West HAC standard errors**: Robust to heteroskedasticity and autocorrelation.
2. **Wild bootstrap** ($B = 1{,}999$, Rademacher weights): Robust to heteroskedasticity with improved finite-sample properties.
3. **Clustered standard errors** (HC1, by event date): Accounts for cross-sectional dependence within events.
4. **Quantile regression** ($\tau = 0.5$): Robust to outliers and heavy tails in the dependent variable.

5. **Permutation test** ($B = 4{,}999$): Gold standard—shuffles semantic labels across events preserving panel structure, providing exact p-values under the null.

All p-values in panel and IRF regressions are adjusted for multiple testing using the Benjamini–Hochberg (BH) procedure within each family of tests (per dependent variable and model type) (Benjamini and Hochberg, 1995). We use only standard significance levels: 1%, 5%, and 10%.

Sub-Period Stability. We examine coefficient stability across six Fed policy regimes: crisis and recovery (2008–2012), post-crisis normalization (2013–2015), pre-pandemic tightening (2016–2019), pandemic response (2020–2021), and inflation tightening (2022–2025).

Alternative Semantic Measures. The dual-model ensemble itself provides a built-in robustness check: by comparing results from MiniLM-only, BERT-only, and ensemble measures, we verify that findings are not driven by model-specific artifacts.

3.3 Hypotheses and Testable Predictions

We now state our hypotheses and, for each one, the coefficient restriction that operationalizes it in the specifications of Section 3.2. Recall the sign conventions: the z-scored stance regressor in equations (28)–(34) is dovish-positive, while the composite MPS in equation (27) is hawkish-positive. Table 1 summarizes the mapping from hypotheses to coefficients, equations, and the tables in which each test is reported.

Hypothesis 1 (Tone Effects on Returns).

- **H1a.** Dovish tone increases returns on risk assets (equities, commodities). *Test:* $\beta_k^{(T)} > 0$ in the semantic decomposition (28) and $\beta_1^{(h)} > 0$ in the local projection (34) for ES and CL.
- **H1b.** Dovish tone decreases returns on safe-haven assets, as accommodative policy triggers portfolio rebalancing away from safe havens toward risk assets. *Test:* $\beta_1^{(h)} < 0$ in (34) for GC, ZN, and ZF.
- **H1c.** Tone effects strengthen with the horizon as markets progressively process policy implications. *Test:* $|\beta_1^{(h)}|$ increasing in h in (34) for ES.

Hypothesis 2 (Novelty Effects on Volatility).

- **H2a.** High novelty increases market volatility through information processing complexity. *Test:* $\beta_2 > 0$ in the event-level regression (33) with ΔRV as the dependent variable, and $\beta > 0$ in the panel regression (31) with Semantic = Novelty, most directly for VX.

- **H2b.** Novelty effects dissipate rapidly as markets resolve uncertainty. *Test:* $|\beta_2^{(h)}|$ decreasing in h in (34) for VX.

- **H2c.** Novelty has minimal effects on directional returns. *Test:* $\beta_2^{(h)} \approx 0$ (statistically indistinguishable from zero) in (34) for the return contracts (ES, CL, GC, DX).

Hypothesis 3 (Policy Surprise Effects).

- **H3a.** Hawkish policy surprises reduce equity returns. *Test:* $\beta_0 < 0$ in the dynamic response function (27) for ES.

- **H3b.** Policy surprises increase implied volatility. *Test:* $\beta_0 > 0$ in (27) for VX.

- **H3c.** Surprise effects on volatility are most pronounced immediately after the announcement. *Test:* $|\beta_k|$ in (27) largest at $k = 0$ and decaying in k for VX.

Hypothesis 4 (Cross-Asset Patterns).

- **H4a.** Risk assets respond to tone in the same direction (risk-on/risk-off). *Test:* $\text{sign}(\beta_1^{(h)})$ equal across ES and CL in (34), opposite for GC, ZN, ZF.

- **H4b.** Volatility responses to novelty are asset-specific. *Test:* β_2 and β_3 in (33) differ in magnitude and significance across the seven contracts.

Finally, the validity of the design itself is testable: under no anticipation of statement content, all pre-announcement coefficients in the placebo specification (35) should be zero, $\beta_1^{(-h)} = \beta_2^{(-h)} = \beta_3^{(-h)} = 0$ for all h.

Table 1: Mapping of Hypotheses to Coefficients, Equations, and Evidence

Hyp.	Prediction	Coefficient restriction (equation)	Evidence
H1a	Dovish tone $\Rightarrow$ risk-asset returns $\uparrow$	$\beta_1^{(h)} > 0$ in (34), ES/CL	Table 9
H1b	Dovish tone $\Rightarrow$ safe-haven returns $\downarrow$	$\beta_1^{(h)} < 0$ in (34), GC/ZN/ZF	Table 9
H1c	Tone effects build with horizon	$\lvert\beta_1^{(h)}\rvert$ increasing in h, ES	Table 9; Fig. 5
H2a	Novelty $\Rightarrow$ volatility $\uparrow$	$\beta_2 > 0$ in (33), ΔRV	Tables 5, 6
H2b	Novelty effects decay quickly	$\lvert\beta_2^{(h)}\rvert$ decreasing in h, VX	Table 10
H2c	Novelty does not move returns	$\beta_2^{(h)} \approx 0$ in (34)	Table 10; Fig. 16
H3a	Hawkish surprise $\Rightarrow$ equity returns $\downarrow$	$\beta_0 < 0$ in (27), ES	Section 4.2
H3b	Surprise $\Rightarrow$ implied volatility $\uparrow$	$\beta_0 > 0$ in (27), VX	Section 4.2
H3c	Surprise effects peak on impact	$\lvert\beta_k\rvert$ max at $k = 0$ in (27), VX	Section 4.2
H4a	Homogeneous return response, risk assets	sign($\beta_1^{(h)}$) equal, ES/CL in (34)	Table 9
H4b	Heterogeneous volatility responses	β_2, β_3 vary across contracts in (33)	Tables 6, 15
Placebo	No pre-announcement effects	$\beta^{(-h)} = 0$ in (35), all contracts	Fig. 8, 18

4 Results

We organize our results as follows. We first describe the semantic variables (Section 4.1), then report baseline responses to the composite MPS (Section 4.2), the tone-novelty decomposition (Section 4.3), and the temporal dynamics of these effects (Section 4.4). Throughout, each hypothesis is evaluated against the coefficient restriction assigned to it in Table 1. The main finding is that tone and novelty load on different dependent variables: tone predicts directional returns, while novelty predicts volatility changes. The VIX model has the highest R^2 (0.78); equity and commodity models follow. Treasury results are weaker.

R^2 values for the return models are low (0.01–0.15), as is typical when predicting minute-level returns where microstructure noise dominates (Barndorff-Nielsen and Shephard, 2002). Low R^2 does not bias the coefficient estimates, but it means that FOMC communication accounts for a small share of total intra-day return variation. The VIX model is the exception: semantic variables explain up to 78% of announcement-window variation in VIX futures, consistent with the hypothesis that these communications operate primarily on uncertainty rather than on the level of returns.

4.1 Descriptive statistics

Table 2 reports descriptive statistics for 1-minute log returns on FOMC days. Table 3 reports summary statistics for the rolling realized measures. Figure 1 plots the average intraday volatility pattern on FOMC days, which spikes sharply at the 14:00 ET release.

Novelty has a mean of 0.065. The scale runs from 0 (a statement identical to the previous one) to 2 (maximally opposite), so a mean of 0.065 corresponds to an average cosine similarity of 0.935 between consecutive statements—most meetings produce only small changes to the prior template. The median is lower (0.025), so the distribution is right-skewed: most meetings produce minor textual changes, but a few produce large departures. The maximum (0.494) occurs during the 2008 crisis, when the Fed introduced zero-lower-bound language (Figure 3). Skewness is 2.45 and kurtosis is 8.16.

The raw tone measure (hawkish-positive) has a mean of 0.309 on a theoretical $[-1, +1]$ scale, where -1 is the dovish centroid, 0 is equidistant, and $+1$ is the hawkish centroid. The positive mean indicates that the average FOMC statement over 2008–2025 lies closer to the hawkish than to the dovish centroid. Because cosine similarities to both centroids are high, realized values occupy a narrow band within the theoretical scale; we therefore z-score tone in all regressions. Skewness is -0.78 and kurtosis is 6.96, driven by the most dovish statements in the sample, which mark its lower extremes (2008: -0.081; 2020: -0.062).

The MPS measure has mean 0.004 and kurtosis 12.18—the fat tails reflect occasional large surprises during economic stress. Jarque-Bera tests reject normality for all three measures ($p < 0.01$), and ARCH-LM tests indicate time-varying variance ($p < 0.05$), which motivates our use of robust standard errors and multiple inference methods throughout.

4.2 Baseline Asset Price Responses to Monetary Policy Surprises

We begin with the composite MPS, testing H3. The dependent variable is the 1-minute log return for each contract; the regressor is the z-scored MPS. A positive coefficient means that hawkish communication surprises are associated with higher returns.

4.2.1 Immediate Market Reactions

The equity market response to hawkish surprises is consistent with H3a (negative equity effect): the ES contract exhibits a -1.2 basis point response ($p < 0.05$). To put this magnitude in perspective, a one-standard-deviation hawkish surprise generates an immediate equity decline roughly equivalent to the average hourly return on a non-announcement day, concentrated in a single minute.

Among commodities, crude oil (CL) falls by 1.0 bps ($p < 0.05$), consistent with tighter policy reducing growth expectations and energy demand. Gold (GC) falls by 0.6 bps, as an inflation hedge becomes less attractive when policy tightens. The Dollar Index (DX) rises by 0.8 bps, consistent with higher expected rate differentials increasing dollar demand.

The VIX response is +11.2 bps ($p < 0.05$, H3b): a one-standard-deviation hawkish surprise is associated with an 11.2 bps increase in VIX futures, about 0.5% of the average VIX level. This is the largest immediate response in our sample. Treasury securities (ZN, ZF) show small, insignificant responses, possibly because our textual MPS is largely orthogonal to the rate expectations already embedded in Treasury futures.

4.2.2 Dynamic Response Evolution

The 45-minute post-announcement window shows that different asset classes adjust at different speeds:

Equities (ES). The initial decline (-1.2 bps at $t = 0$) deepens to -3.8 bps by $t = 15$ and -6.8 bps by $t = 45$. The cumulative response is 5.7 times the instantaneous reaction, so the first-minute price change captures only a fraction of the total adjustment.

VIX futures. The initial spike (+11.2 bps at $t = 0$) loses roughly half its magnitude by $t = 20$ and returns near its pre-announcement level by $t = 45$.

Treasuries (ZN, ZF). Responses become marginally significant only at $t =$ 30–45, consistent with slower transmission to term premiums.

Commodities (CL, GC). Crude oil effects continue to grow through $t = 45$, consistent with the slower adjustment typical of physical commodity markets.

These patterns support H3c for volatility—the VIX response peaks on impact and decays—but equity and commodity responses continue to build, foreshadowing the gradual tone effects documented below. The speed of adjustment varies across assets, which motivates the asset-specific decomposition.

4.3 Semantic Decomposition: Tone vs. Novelty Effects

We now separate tone from novelty. The dependent variables are log returns (testing H1) and realized volatility changes (testing H2). The regressors are the z-scored tone and novelty measures, entered separately and jointly.

4.3.1 Tone Effects on Asset Returns (H1)

A one-standard-deviation dovish shift in tone is associated with ES gains that build steadily over the post-announcement window: 3.8 bps by $h = 30$, 6.8 bps by $h = 60$, and 12.0 bps by $h = 120$ (Table 9, H1a)—several days of average equity returns compressed into two hours. This is consistent with lower discount rates, higher growth expectations, and greater risk appetite.

Safe-haven assets move in the opposite direction: at $h = 15$, gold falls by 3.1 bps and 10-year Treasuries (ZN) by 1.1 bps following dovish tone (H1b), consistent with a "risk-on" rotation out of safe havens. The Dollar Index response is small and changes sign across horizons (+1.4 bps at $h = 15$, -4.0 bps at $h = 120$).

Tone effects grow stronger, not weaker, over the post-announcement window: the ES coefficient rises monotonically from 0.2 bps at $h = 5$ to 3.8 bps at $h = 30$ and 12.0 bps at $h = 120$ (Table 9, H1c). This gradual amplification—rather than the immediate level shift that a frictionless model would imply—is consistent with sequential portfolio adjustment by heterogeneous participants (algorithms, institutions, retail).

4.3.2 Novelty Effects on Volatility (H2)

Novelty loads on volatility, not on returns. The VIX response to a one-standard-deviation increase in novelty is -5.57 bps ($p < 0.05$), meaning that genuinely new language is associated with *lower* implied volatility. This sign is opposite to our ex ante prediction (H2a) but has a straightforward interpretation: a statement that departs substantially from the prior meeting's language sends a clearer signal about the Fed's current assessment, helping investors narrow the range of possible outcomes. A repetitive statement, by contrast, leaves open the question of whether the unchanged wording reflects genuine stability or simply a failure to update the language.

The VIX novelty coefficient is largest in the first 5 minutes and falls to near zero by $t = 30$ (H2b). Novelty coefficients on directional returns (ES, CL, GC, DX) are small and insignificant (H2c). The combination—novelty predicts volatility but not returns—separates our measure from standard "surprise" variables, which conflate direction and information content.

4.3.3 Panel Minute-Level Results

Tables 4–5 report panel regressions of log(RV) on the interaction of post-announcement indicators with semantic measures. The coefficient on Post×Stance is negative and significant at 1% for 6 of 7 contracts: ES (−5.31), VX (−3.28), ZN (−5.25), ZF (−4.16), CL (−5.87), and GC (−6.03). Only DX is insignificant (+0.15). The negative sign means that dovish stance is associated with lower post-announcement realized volatility—accommodative statements calm markets—while hawkish statements are followed by larger volatility increases, consistent with contractionary signals generating more repricing. By contrast, the Post×Novelty coefficients (Table 5) are smaller and significant only for the Treasury contracts (ZN: −1.81, $p < 0.01$; ZF: −1.22, $p < 0.10$).

4.3.4 Event-Level Results

Tables 6–8 report event-level regressions with the interaction specification $\Delta Y_i = \alpha + \beta_1 \text{Stance} + \beta_2 \text{Novelty} + \beta_3 (\text{Stance} \times \text{Novelty}) + \varepsilon_i$. The stance×novelty interaction is negative and significant for VIX (−7.96, $t = -3.12$, $p < 0.01$), ZF (−2.48, $t = -2.22$), ZN (−1.66, $t = -1.99$), and CL (−8.61, $p < 0.10$); Figure 4 summarizes the coefficient estimates as a heatmap. The interaction indicates that stance and novelty reinforce each other: the volatility decline associated with a dovish statement is larger when the statement also departs substantially from prior language, whereas a statement that repeats the previous meeting's wording has a smaller volatility effect regardless of its tone.

R^2 values are much higher for volatility than for returns: ΔRV regressions achieve 0.286 (ES), 0.268 (CL), and 0.245 (ZF), versus below 0.05 for most return regressions. The semantic variables explain more of the variation in volatility (which depends on information content) than in returns (which depend on the direction of the surprise in a noisy environment).

4.3.5 Local Projection Impulse Response Functions

Tables 9–14 and Figures 6–7 report local projection estimates. The VIX response to stance is negative at every horizon, statistically strongest between 5 and 30 minutes (-14.74 bps at $h = 5$, -22.03 at $h = 30$), and attenuates thereafter (-14.27 at $h = 120$, with much wider standard errors); the profile is hump-shaped, peaking near 30 minutes. The stance×novelty interaction on abnormal VIX returns (Figure 7) is significant from $h = 5$ through $h = 120$, with the t-statistic reaching -5.06.

The ES coefficient on stance grows monotonically from near zero at $h = 5$ to 12.0 bps at $h = 120$ (standard errors also grow). A frictionless model would predict an immediate level shift; the gradual increase we observe is more consistent with sequential adjustment by heterogeneous investors.

4.3.6 Cross-Asset Patterns (H4)

Risk assets (ES, CL) both rise after dovish tone, with similar timing (H4a). Safe havens (GC, ZN, ZF) fall, consistent with risk-on/risk-off. But the volatility responses differ across assets (H4b): VIX responds most to the stance×novelty interaction, Treasury volatility responds mainly to stance alone, and commodity volatility responds to both. The return homogeneity and volatility heterogeneity suggest that the two channels operate with different relative strength across asset classes.

4.3.7 Pre-Announcement Placebo Tests

Figures 8–18 report pre-announcement placebo tests: we estimate the same IRF specification for cumulative returns from $-h$ to 0. If statement content is not anticipated, all pre-announcement coefficients should be zero. None is significant at 10% across any ticker, horizon, or semantic measure. This rules out information leakage and pre-existing trends.

4.3.8 Multi-Method Robustness

Tables 15–16 report five alternative inference methods for every event-level coefficient. We call a result "robust" if at least 3 of 5 methods give $p < 0.10$. The most robust results are: stance on $\Delta \log \mathrm{RV}$ for ES, CL, GC, ZF, and ZN (each #Sig = 3); stance on ΔRV for ES, CL, and DX (each #Sig = 3); and novelty on ΔRV for ES and CL (#Sig = 4). Figure 9 displays these counts. Sub-period analysis (Figure 19) shows consistent coefficient signs across six Fed regimes, though magnitudes vary.

With 7 assets and multiple dependent variables, we test many coefficients; results with #Sig = 1 should be treated with caution. We rely on the Benjamini–Hochberg correction and the robustness counts to limit false discovery.

4.4 Economic Interpretation: Temporal Dynamics and Half-Lives

We summarize the speed of adjustment using half-lives: the time for an initial effect to fall to half its peak. For exponential decay, $t_{1/2} = \ln(2)/\lambda$.

The VIX response (+11.2 bps on impact) falls to half its peak within 15–20 minutes; reading the decay as approximately exponential gives a half-life of about 17.5 minutes. Once the statement text is known, uncertainty about the Fed's message dissipates quickly.

Equity returns show the opposite pattern. The ES response to dovish tone *grows* steadily over the post-announcement window, roughly tripling between $h = 30$ (3.8 bps) and $h = 120$ (12.0 bps). Fundamental repricing—revising expected cash flows and discount rates—takes longer than uncertainty resolution, consistent with gradual diffusion across heterogeneous investors.

The two patterns together—fast volatility decay and slow return amplification—are hard to reconcile with a single channel. The informational dimension (is the statement new?) resolves quickly; the directional dimension (is it hawkish or dovish?) takes longer to price in.

5 Conclusion

5.1 Summary of Findings

We summarize our findings under two headings: what we learn about the economics of monetary policy transmission, and what we contribute methodologically.

Tone predicts directional returns. Dovish tone is associated with higher equity returns that build to about 12 bps over two hours, and with lower safe-haven returns (gold: −3.1 bps; Treasuries: −1.1 bps at 15 minutes), supporting H1a–b. This is consistent with the standard transmission mechanism in Bernanke and Kuttner (2005) and Gürkaynak et al. (2005)—accommodative expectations lower discount rates and shift portfolios toward risk assets—but we show that the channel operates through the *qualitative* content of the statement, not only through the rate decision.

Tone effects grow stronger over time: the ES coefficient rises from 0.2 bps at $h = 5$ to 12.0 bps at $h = 120$ (H1c). Rosa (2013) documents continued price adjustment over hours after FOMC releases; our minute-level data show that this amplification is consistent with heterogeneous processing speeds across market participants (Veldkamp, 2011).

Novelty predicts lower volatility, not higher. Novelty is associated with *reduced* implied volatility (VIX: -5.57 bps, $p < 0.05$; interaction $t = -5.06$), which reverses the sign we predicted under H2a. Prior work on disagreement-driven volatility (Patton and Verardo, 2012) and news-based uncertainty (Manela and Moreira, 2017) would predict the opposite. The difference, we believe, is that in the specific setting of FOMC statements, a departure from prior language acts as a *signal of clarity*: the Fed is actively updating its message, which narrows the set of plausible policy paths. A repetitive statement, conversely, leaves open whether the Fed's views have changed but the language has not been updated. Novelty effects dissipate in under 10 minutes (H2b) and do not predict returns (H2c), confirming that this channel operates on volatility rather than on prices.

The composite surprise replicates known patterns. Hawkish surprises are associated with 1–7 bps lower equity returns (H3a) and 11.2 bps higher VIX (H3b), with volatility effects concentrated in the first minutes (H3c). That our textual MPS produces results consistent with the rate-surprise literature (Kuttner, 2001; Bernanke and Kuttner, 2005; Gürkaynak et al., 2005; Savor and Wilson, 2014) reassures us that the NLP pipeline captures economically relevant variation.

Cross-asset patterns differ between returns and volatility. Risk assets (ES, CL) respond to tone with the same sign and similar timing (H4a). Safe havens (GC, ZN, ZF) move in the opposite direction. But volatility responses vary: VIX loads on the interaction, Treasury volatility on stance alone, commodity volatility on both. This pattern is consistent with Fleming and Remolona (1999) and Balduzzi et al. (2001), who document heterogeneous announcement effects across asset classes.

Volatility resolves fast; returns adjust slowly. The VIX half-life is about 17.5 minutes; the ES tone effect continues to build for two hours after the announcement. A single-channel model cannot produce both patterns simultaneously. The fast channel (novelty → volatility) reflects the resolution of uncertainty about what the Fed will say. The slow channel (tone → returns) reflects the time needed to revise discount rates, growth expectations, and portfolio allocations across a heterogeneous investor base.

5.2 Methodological Contributions

On the measurement side, we make three contributions. First, using two architecturally different models (MiniLM, 33M parameters; BERT, 110M parameters) and comparing their outputs reduces the risk that any finding is an artifact of one particular architecture. The inter-model confidence (mean 0.837 for novelty) flags statements where the two models disagree. Second, two-stage domain adaptation (TSDAE then MNRL) yields axis separations of 0.248–0.302, well above the 0.04 minimum. Third, PCA-based reference selection replaces subjective date choices with an algorithm, removing a source of researcher discretion.

Working at the minute level (148 events $\times$ $\pm$120 minutes) matters for two reasons: it allows causal identification from the pre-determined release time (validated by the placebo tests), and it reveals the fast-volatility / slow-return asymmetry that daily data would average away.

5.3 Implications for Policy and Communication Strategy

Three findings have implications for how the Fed drafts its statements.

First, novelty reduces volatility. The conventional view is that central banks should change their language gradually to avoid "surprising" markets. Our estimates say the opposite: substantive changes in wording are associated with *lower* implied volatility, presumably because a clear departure from the prior statement narrows the range of plausible interpretations. When the Fed needs to signal a regime change—a new framework, a pivot from tightening to easing—a decisive rewrite of the statement may be less destabilizing than incremental edits.

Second, tone has persistent valuation effects. Because the equity response to tone grows over 2 hours rather than being absorbed immediately, the hawkish-dovish framing of the statement has real wealth consequences: a one-standard-deviation dovish shift is associated with up to 12 bps of equity returns, which, on a market capitalization of roughly $40 trillion, is economically meaningful.

Third, the interaction matters. The stance$\times$novelty interaction on VIX ($t = -5.06$, persistent from 5 to 120 minutes) implies that a directionally clear statement packaged in new language reduces uncertainty more than the same message delivered in boilerplate.

5.4 Limitations

We note several limitations.

Sample period. Our 2008–2025 sample is dominated by extraordinary monetary policy regimes: the zero lower bound (2008–2015), quantitative easing, and the post-pandemic tightening cycle. While our sub-period stability analysis (Figure 19) shows that coefficient signs are generally consistent across six Fed policy regimes, we cannot rule out that the magnitude of communication effects differs in more "normal" policy environments. In particular, the ZLB period may overstate the importance of qualitative communication (since rate decisions were constrained, statements became the primary policy instrument), while the recent tightening cycle may understate novelty effects (since rate hikes were widely anticipated, reducing the scope for genuine communication surprises).

Measurement error. Our NLP-based tone and novelty measures are subject to measurement error whose properties are difficult to characterize fully. Classical measurement error would attenuate our coefficient estimates toward zero (creating an "errors-in-variables" bias), making our significant findings conservative. However, if measurement error is correlated with meeting characteristics (e.g., if our models systematically misclassify statements from certain Fed Chairs or policy regimes), the bias could go in either direction. The inter-model agreement diagnostics (mean confidence 0.837 for novelty, 0.562–0.806 for tone) help identify statements where measurement may be less reliable, but they do not fully resolve this concern.

Multiple testing. With 7 assets, 5 dependent variables, and multiple inference methods, we test a large number of coefficients. We address this through Benjamini–Hochberg FDR correction within each family of tests and by reporting multi-method robustness counts. Nevertheless, some individually significant results—particularly those with #Sig = 1 in the robustness tables—should be interpreted cautiously as potentially reflecting false discovery.

Identification threats. Our identification strategy exploits the pre-determined timing and content of FOMC statements, validated by pre-announcement placebo tests. However, we cannot fully rule out confounding from simultaneous information releases. FOMC announcements sometimes coincide with the Summary of Economic Projections ("dot plot"), which provides quantitative rate path forecasts that may interact with our textual measures. Similarly, market expectations about the subsequent press conference (beginning 30 minutes after the statement) may influence post-announcement price dynamics within our event window. These concurrent information flows are not fully separable from the statement text itself.

Generalizability. We study U.S. futures markets around FOMC statements only. Whether the same tone-return and novelty-volatility patterns hold for other central banks (ECB, BOJ, BOE), other Fed communications (minutes, speeches), or other asset classes (corporate bonds, emerging-market equities) is an open question.

5.5 Future Directions

Several extensions seem natural.

More semantic dimensions. We measure tone and novelty; FOMC statements also vary in uncertainty language, temporal focus (forward- vs. backward-looking), and internal consensus (voting dissents). The PCA axis construction can accommodate additional dimensions.

International markets. FOMC announcements move global asset prices (Wongswan, 2009; Ehrmann et al., 2011). Applying the tone-novelty decomposition to international data would show whether both channels transmit across borders or whether one dominates.

Structural models. Our evidence is reduced-form. A model with heterogeneous agents and differential processing of directional versus informational content could rationalize the fast-volatility / slow-return asymmetry and generate further predictions.

Joint analysis with rate surprises. Our textual MPS and the rate-based surprise of Kuttner (2001) measure different things. Estimating both jointly would reveal whether they contain complementary information and whether the information-versus-policy decomposition of Jaroćinski and Karádi (2020); Nakamura and Steinsson (2018) maps onto our tone-versus-novelty decomposition.

Tables

Table 2: Descriptive Statistics: 1-Minute Log Returns on FOMC Days

Ticker	N	Mean	SD	Skew	Kurt	JB	Min	Median	Max	AC(1)
E-mini S&P 500	42,180	-0.030	3.70	0.96	208.0	73858777***	-118.7	0.000	148.0	-0.022
VIX Futures	37,050	0.072	14.12	0.02	77.9	8665376***	-466.1	0.000	419.8	-0.077
10Y T-Note	42,180	0.009	1.07	1.21	44.4	3027292***	-13.8	0.000	30.8	-0.103
5Y T-Note	42,180	0.004	0.68	-4.70	444.1	342093523***	-43.8	0.000	17.6	-0.059
Dollar Index	42,180	-0.000	1.69	-1.74	1508.6	3984214777***	-119.2	0.000	115.4	0.194
Crude Oil WTI	42,180	-0.006	5.14	-0.64	166.7	47098526***	-191.8	0.000	143.9	-0.062
Gold	41,895	0.001	2.91	0.24	121.3	24424589***	-91.5	0.000	108.8	-0.017

Notes: All statistics in basis points except N, AC(1), and JB. Sample includes all 1-minute observations on FOMC event days ± 1 day with grid regularization and forward-filling of short gaps. Skewness and excess kurtosis are sample moments. JB is the Jarque–Bera test statistic. AC(1) denotes the first-order autocorrelation coefficient. Negative AC(1) values reflect bid–ask bounce effects typical of high-frequency data (Andersen and Bollerslev, 1997). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 3: Descriptive Statistics: Rolling Realized Measures (5-Minute Window)

	N	Realized Volatility (bps)					log(RV)		Realized Beta vs. ES		
Contract		Mean	SD	P5	P50	P95	Mean	SD	Mean	SD	P50
E-mini S&P 500	35,668	3.43	5.68	0.00	2.41	10.38	-11.78	6.55	1.000	0.000	1.000
10Y T-Note	35,668	1.41	1.63	0.00	1.31	3.81	-14.11	7.09	-0.018	0.395	0.000
5Y T-Note	35,668	0.79	1.15	0.00	0.66	2.36	-14.77	6.86	-0.007	0.225	0.000
Dollar Index	35,668	0.78	3.45	0.00	0.00	3.53	-19.32	6.34	-0.010	0.482	0.000
Crude Oil WTI	35,668	5.86	9.47	0.00	3.97	17.90	-11.24	6.67	0.184	1.472	0.006
Gold	35,427	3.64	4.64	0.00	2.73	10.51	-11.29	6.30	0.044	0.958	0.000
VIX Futures	31,330	11.16	23.33	0.00	0.00	52.40	-18.31	7.74	-0.714	4.451	0.000

Notes: Statistics computed across all minute-level observations within ±30 minutes of FOMC announcements ($K = 5$ min rolling window, NA-tolerant with up to 20% missing data). RV is the NA-tolerant rolling realized variance defined in Section 3, displayed in square-root (volatility) units in basis points; log(RV) is computed on the realized variance in raw decimal-return units, so its level is not directly comparable to the bps columns. $\beta_t^{\mathrm{real}} = \mathrm{RCov}(r_i, r_{\mathrm{ES}})/\mathrm{RVar}(r_{\mathrm{ES}})$. ES beta is 1.000 by construction. log(RV) is near-Gaussian, validating its use as a regression dependent variable.

Table 4: Panel Regressions: log(RV) on Post × Stance (5-min, ±30 min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-14.8474***	-19.9268***	-17.7088***	-17.9767***	-22.6514***	-14.9591***	-15.0663***
	(0.4794)	(0.4462)	(0.4275)	(0.4466)	(0.1175)	(0.4704)	(0.4845)
Post × Stance	-5.3084***	-3.2805***	-5.2530***	-4.1583***	0.1481	-5.8682***	-6.0348***
	(0.8271)	(0.7803)	(0.6490)	(0.7052)	(0.2836)	(0.8526)	(0.8609)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.134	0.047	0.133	0.091	0.001	0.149	0.166
Adj. R^2	0.134	0.047	0.132	0.091	0.001	0.149	0.166

Notes: Dependent variable: $\log(\text{RV}_t(5))$. Post×Stance $= \mathbf{1}[t > 0] \times$ Stance (z-scored). Post×Novelty $= \mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ±30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 5: Panel Regressions: log(RV) on Post × Novelty (5-min, ±30 min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-13.4894***	-18.9261***	-16.4259***	-16.9510***	-22.6505***	-13.4467***	-13.4834***
	(0.2892)	(0.3813)	(0.3396)	(0.3423)	(0.1015)	(0.2887)	(0.2946)
Post × Novelty	-0.5169	0.3020	-1.8142***	-1.2188*	0.8439	-0.3328	-0.4821
	(0.5996)	(0.8881)	(0.5683)	(0.5493)	(0.4881)	(0.6298)	(0.6316)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.003	0.001	0.031	0.015	0.069	0.001	0.002
Adj. R^2	0.002	0.001	0.031	0.015	0.068	0.001	0.002

Notes: Dependent variable: $\log(\mathrm{RV}_t(5))$. Post×Stance $= \mathbf{1}[t > 0] \times$ Stance (z-scored). Post×Novelty $= \mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ±30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 6: Rolling Delta RV Regressions (30-min Window, Interaction Model)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	5.54***	26.00***	2.23***	1.65***	-0.40	12.32***	9.96***
	(1.67)	(6.51)	(0.64)	(0.49)	(0.62)	(2.02)	(1.51)
Stance	-4.94**	-2.80	-1.50*	-1.30*	-1.65	-7.78**	-1.05
	(1.53)	(6.03)	(0.71)	(0.64)	(0.90)	(2.65)	(1.92)
Novelty	6.82*	3.42	0.70	0.94	1.15	8.20*	3.34*
	(3.42)	(5.98)	(0.81)	(0.83)	(1.09)	(3.63)	(1.35)
Stance × Novelty	-5.32	-7.96**	-1.66*	-2.48*	-1.20	-8.61*	-4.68
	(3.15)	(2.55)	(0.83)	(1.12)	(1.39)	(3.60)	(3.13)
N	148	130	148	148	148	148	147
R^2	0.286	0.054	0.131	0.245	0.137	0.268	0.128
Adj. R^2	0.271	0.032	0.113	0.229	0.119	0.253	0.109

Notes: Dependent variable: Delta RV (bps) (Rolling 30-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Coefficients $\times 10,000$. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 7: Rolling Delta log(RV) Regressions (30-min Window, Interaction Model)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	6.1744***	6.8238***	7.1287***	6.9380***	-0.3664	8.4210***	8.4805***
	(1.0374)	(1.6303)	(1.0670)	(1.0305)	(0.3671)	(0.8868)	(0.8682)
Stance	-2.9143	-0.8191	-2.5521	-2.3930	-0.9082	-2.1849	-1.7794
	(1.0198)	(1.8314)	(1.1059)	(1.0661)	(0.5066)	(1.0687)	(1.0675)
Novelty	0.3441	-0.2304	-0.7730	-0.6468	0.5076	0.8902	0.7810
	(0.6037)	(0.9547)	(0.5429)	(0.5328)	(0.5352)	(0.4145)	(0.3970)
Stance × Novelty	-0.6767	-1.5107	-1.0145	-1.0850	-0.3540	-0.7406	-0.6972
	(0.6652)	(0.7161)	(0.5278)	(0.5025)	(0.7771)	(0.6279)	(0.6369)
N	148	130	148	148	148	148	147
R^2	0.064	0.018	0.075	0.071	0.055	0.052	0.041
Adj. R^2	0.044	-0.005	0.056	0.051	0.036	0.033	0.021

Notes: Dependent variable: $\Delta \log(\text{RV})$ (Rolling 30-min). Coefficients are in log-units; a coefficient of -2.91 on Stance for ES means that a one-standard-deviation hawkish shift is associated with a 2.91 log-unit decrease in realized volatility. All regressors z-scored. Clustered SE in parentheses. Stars: BH-adjusted p-values. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 8: Rolling Realized Beta Change Regressions (30-min Window)

	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-0.2364	0.0098	-0.0025	0.0058	-0.0933	0.1275*
	(0.1651)	(0.0192)	(0.0134)	(0.0074)	(0.0650)	(0.0758)
Stance	-0.0190	0.0118	0.0195	0.0137	-0.0008	0.1244
	(0.2653)	(0.0198)	(0.0158)	(0.0093)	(0.0891)	(0.0869)
Novelty	0.1096	-0.0096	-0.0064	-0.0086	-0.1290	-0.0539
	(0.0917)	(0.0102)	(0.0073)	(0.0075)	(0.1020)	(0.0344)
Stance × Novelty	-0.0818	0.0064	0.0074	0.0011	-0.1772	-0.0852
	(0.1018)	(0.0118)	(0.0074)	(0.0080)	(0.2106)	(0.0755)
N	71	88	88	88	88	87
R^2	0.009	0.003	0.008	0.045	0.064	0.047
Adj. R^2	-0.036	-0.033	-0.027	0.010	0.031	0.013

Notes: Dependent variable: $\Delta\hat{\beta}$ (change in rolling realized beta relative to ES, 30-min window). Coefficients are in beta units; a coefficient of 0.12 means that a one-standard-deviation increase in the regressor is associated with a 0.12 increase in co-movement with the S&P 500. All regressors z-scored. Clustered SE in parentheses. Stars: BH-adjusted p-values. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 9: Price IRF: Policy Stance on Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	0.23	1.98	3.76	6.82	11.98
	(0.92)	(2.11)	(2.76)	(3.31)	(6.76)
VIX Futures	-14.74	-19.11	-22.03	-17.73	-14.27
	(6.26)	(8.33)	(9.47)	(14.62)	(31.57)
10Y T-Note	-0.61	-1.14	-0.90	-1.47	-3.88
	(0.43)	(0.85)	(0.93)	(0.92)	(2.20)
5Y T-Note	0.89	0.63	0.90	0.79	-0.71
	(0.34)	(0.35)	(0.55)	(0.51)	(1.34)
Dollar Index	1.01	1.43	1.05	-0.45	-4.01
	(0.43)	(0.61)	(0.45)	(0.89)	(2.14)
Crude Oil WTI	-2.10	-0.81	4.75	3.06	9.32
	(4.28)	(6.28)	(8.46)	(10.23)	(9.69)
Gold	-1.47	-3.08	0.51	0.62	1.31
	(1.30)	(1.42)	(2.04)	(2.46)	(3.96)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1\text{Stance} + \beta_2\text{Novelty} + \beta_3\text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Policy Stance reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. Pre-announcement placebo test passed: no significant pre-event coefficients.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 10: Price IRF: Novelty on Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	-0.99	-1.42	-0.65	-0.61	-3.91
	(0.56)	(1.70)	(1.34)	(1.82)	(2.25)
VIX Futures	10.24	4.84	-2.55	-6.02	6.22
	(9.19)	(9.31)	(7.62)	(5.36)	(8.33)
10Y T-Note	0.82	1.17	1.52	1.87	3.70
	(0.60)	(0.97)	(0.91)	(1.09)	(2.15)
5Y T-Note	-0.61	-0.16	0.16	-0.00	1.08
	(0.41)	(0.10)	(0.17)	(0.27)	(0.47)
Dollar Index	0.41	0.30	0.61	0.94	0.60
	(1.34)	(1.74)	(1.51)	(1.62)	(2.22)
Crude Oil WTI	-0.52	-0.82	-3.78	-2.03	-1.29
	(3.54)	(4.29)	(7.83)	(6.83)	(6.20)
Gold	-1.30	0.84	0.55	-1.13	2.00
	(1.42)	(2.62)	(2.33)	(1.89)	(2.36)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1 \text{Stance} + \beta_2 \text{Novelty} + \beta_3 \text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Novelty reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. Pre-announcement placebo test passed: no significant pre-event coefficients.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 11: Price IRF: Stance × Novelty on Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	2.78	6.91	8.27	9.47	6.35
	(1.36)	(4.14)	(4.39)	(4.61)	(4.49)
VIX Futures	-28.02**	-47.97**	-61.74***	-51.80***	-44.06*
	(8.62)	(12.75)	(15.54)	(10.66)	(14.53)
10Y T-Note	-1.76	-3.53	-3.68	-4.77	-5.70
	(0.85)	(1.68)	(1.74)	(1.88)	(2.84)
5Y T-Note	1.15	-0.03	0.24	-0.19	0.63
	(0.57)	(0.10)	(0.16)	(0.28)	(0.46)
Dollar Index	2.96	4.05	3.16	3.32	2.83
	(1.15)	(1.53)	(1.27)	(1.41)	(2.21)
Crude Oil WTI	7.62	7.77	19.17	16.07	18.35
	(7.45)	(11.46)	(18.20)	(19.84)	(18.66)
Gold	1.48	-5.97*	-2.19	1.94	-0.60
	(1.59)	(2.08)	(1.57)	(2.34)	(2.13)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1\text{Stance} + \beta_2\text{Novelty} + \beta_3\text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Stance × Novelty reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. Pre-announcement placebo test passed: no significant pre-event coefficients.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 12: Price IRF: Policy Stance on Abnormal Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	–	–	–	–	–
VIX Futures	-15.04	-22.74	-28.86	-28.65	-30.54
	(7.07)	(9.76)	(10.84)	(16.52)	(37.21)
10Y T-Note	-0.85	-3.12	-4.67	-8.29	-15.86
	(1.20)	(2.81)	(3.59)	(4.03)	(8.44)
5Y T-Note	0.66	-1.35	-2.86	-6.03	-12.68
	(0.85)	(2.20)	(3.10)	(3.63)	(7.84)
Dollar Index	0.78	-0.55	-2.72	-7.27	-15.99
	(0.92)	(1.86)	(2.63)	(3.73)	(8.39)
Crude Oil WTI	-2.33	-2.79	0.99	-3.76	-2.65
	(3.71)	(5.05)	(7.17)	(8.73)	(8.44)
Gold	-1.77	-5.75	-5.34	-8.84**	-13.76
	(1.25)	(2.57)	(2.54)	(2.66)	(5.73)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1\text{Stance} + \beta_2\text{Novelty} + \beta_3\text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Policy Stance reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. ES entries are omitted because abnormal returns are defined relative to the ES benchmark and are zero by construction. Pre-announcement placebo test passed: no significant pre-event coefficients.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 13: Price IRF: Novelty on Abnormal Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	–	–	–	–	–
VIX Futures	10.21	1.84	-4.80	-10.09	7.18
	(9.77)	(10.47)	(8.32)	(6.08)	(10.24)
10Y T-Note	1.81	2.59	2.17	2.48	7.61
	(0.97)	(2.14)	(1.87)	(2.07)	(3.41)
5Y T-Note	0.38	1.26	0.81	0.61	5.00
	(0.58)	(1.71)	(1.37)	(1.91)	(2.57)
Dollar Index	1.40	1.72	1.26	1.56	4.51
	(1.48)	(2.77)	(2.28)	(2.92)	(3.68)
Crude Oil WTI	0.47	0.60	-3.13	-1.42	2.62
	(3.34)	(3.64)	(7.50)	(6.75)	(5.73)
Gold	-0.30	2.34	1.43	-0.21	6.27
	(1.46)	(3.08)	(2.70)	(2.63)	(2.70)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1\text{Stance} + \beta_2\text{Novelty} + \beta_3\text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Novelty reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. ES entries are omitted because abnormal returns are defined relative to the ES benchmark and are zero by construction. Pre-announcement placebo test passed: no significant pre-event coefficients.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 14: Price IRF: Stance × Novelty on Abnormal Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	–	–	–	–	–
VIX Futures	-31.74**	-58.80***	-72.89***	-65.76***	-53.38*
	(9.50)	(14.98)	(18.63)	(13.00)	(17.64)
10Y T-Note	-4.54	-10.44	-11.95	-14.24	-12.05
	(2.09)	(5.70)	(6.03)	(6.30)	(6.89)
5Y T-Note	-1.63	-6.94	-8.03	-9.66	-5.72
	(0.99)	(4.18)	(4.36)	(4.58)	(4.66)
Dollar Index	0.18	-2.85	-5.11	-6.15	-3.52
	(1.65)	(3.98)	(4.37)	(5.05)	(4.70)
Crude Oil WTI	4.84	0.86	10.90	6.59	12.00
	(6.44)	(8.45)	(14.43)	(16.04)	(15.12)
Gold	-1.30	-12.83	-10.31	-7.35	-6.73
	(1.32)	(5.31)	(4.61)	(3.60)	(4.95)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1 \text{Stance} + \beta_2 \text{Novelty} + \beta_3 \text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Stance × Novelty reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. ES entries are omitted because abnormal returns are defined relative to the ES benchmark and are zero by construction. Pre-announcement placebo test passed: no significant pre-event coefficients.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 15: Robustness: Event-Level Regressions — ΔRV and Δ logRV (30-Minute Window, Interaction Model)

Contract	Dep. Var.	Regressor	$\hat{\beta}$	t	p_{NW}	p_{Boot}	p_{Clust}	p_{QReg}	p_{Perm}	#Sig
Crude Oil WTI	ΔRV	Stance	-7.78	-2.94	0.004	0.691	0.001	–	0.007	3
Crude Oil WTI	ΔRV	Novelty	8.20	2.26	0.025	0.068	0.009	–	0.000	4
Crude Oil WTI	ΔRV	St. × Nov.	-8.61	-2.39	0.018	0.212	0.014	–	0.002	3
Dollar Index	ΔRV	Stance	-1.65	-1.83	0.069	0.260	0.095	–	0.014	3
Dollar Index	ΔRV	Novelty	1.15	1.06	0.293	0.069	0.366	–	0.014	2
Dollar Index	ΔRV	St. × Nov.	-1.20	-0.86	0.392	0.124	0.469	–	0.047	1
E-mini S&P 500	ΔRV	Stance	-4.94	-3.22	0.002	0.584	0.011	–	0.019	3
E-mini S&P 500	ΔRV	Novelty	6.82	2.00	0.048	0.013	0.028	–	0.002	4
E-mini S&P 500	ΔRV	St. × Nov.	-5.32	-1.69	0.094	0.150	0.167	–	0.024	2
Gold	ΔRV	Stance	-1.05	-0.55	0.583	0.692	0.602	–	0.561	0
Gold	ΔRV	Novelty	3.34	2.48	0.014	0.318	0.056	–	0.021	3
Gold	ΔRV	St. × Nov.	-4.68	-1.49	0.138	0.266	0.117	–	0.025	1
VIX Futures	ΔRV	Stance	-2.80	-0.46	0.644	0.659	0.580	–	0.619	0
VIX Futures	ΔRV	Novelty	3.42	0.57	0.568	0.492	0.486	–	0.382	0
VIX Futures	ΔRV	St. × Nov.	-7.96	-3.12	0.002	0.601	0.058	–	0.067	3
5Y T-Note	ΔRV	Stance	-1.30	-2.04	0.043	0.236	0.235	–	0.050	2
5Y T-Note	ΔRV	Novelty	0.94	1.13	0.260	0.203	0.235	–	0.041	1
5Y T-Note	ΔRV	St. × Nov.	-2.48	-2.22	0.028	0.070	0.149	–	0.004	3
10Y T-Note	ΔRV	Stance	-1.50	-2.10	0.038	0.210	0.210	–	0.025	2
10Y T-Note	ΔRV	Novelty	0.70	0.87	0.388	0.314	0.313	–	0.082	1
10Y T-Note	ΔRV	St. × Nov.	-1.66	-1.99	0.048	0.135	0.250	–	0.013	2
Crude Oil WTI	Δ logRV	Stance	-2.1849	-2.04	0.043	0.442	0.035	–	0.022	3
Crude Oil WTI	Δ logRV	Novelty	0.8902	2.15	0.033	0.614	0.086	–	0.135	2
Crude Oil WTI	Δ logRV	St. × Nov.	-0.7406	-1.18	0.240	0.650	0.203	–	0.326	0
Dollar Index	Δ logRV	Stance	-0.9082	-1.79	0.075	0.337	0.121	–	0.041	2
Dollar Index	Δ logRV	Novelty	0.5076	0.95	0.345	0.209	0.478	–	0.064	1
Dollar Index	Δ logRV	St. × Nov.	-0.3540	-0.46	0.649	0.464	0.673	–	0.247	0
E-mini S&P 500	Δ logRV	Stance	-2.9143	-2.86	0.005	0.716	0.001	–	0.004	3
E-mini S&P 500	Δ logRV	Novelty	0.3441	0.57	0.570	0.639	0.609	–	0.577	0
E-mini S&P 500	Δ logRV	St. × Nov.	-0.6767	-1.02	0.311	0.644	0.249	–	0.389	0
Gold	Δ logRV	Stance	-1.7794	-1.67	0.098	0.433	0.093	–	0.066	3
Gold	Δ logRV	Novelty	0.7810	1.97	0.051	0.583	0.120	–	0.168	1
Gold	Δ logRV	St. × Nov.	-0.6972	-1.09	0.275	0.618	0.236	–	0.334	0
VIX Futures	Δ logRV	Stance	-0.8191	-0.45	0.655	0.610	0.563	–	0.558	0
VIX Futures	Δ logRV	Novelty	-0.2304	-0.24	0.810	0.816	0.794	–	0.817	0
VIX Futures	Δ logRV	St. × Nov.	-1.5107	-2.11	0.037	0.781	0.037	–	0.192	2
5Y T-Note	Δ logRV	Stance	-2.3930	-2.24	0.026	0.545	0.012	–	0.014	3
5Y T-Note	Δ logRV	Novelty	-0.6468	-1.21	0.227	0.624	0.164	–	0.293	0
5Y T-Note	Δ logRV	St. × Nov.	-1.0850	-2.16	0.032	0.827	0.024	–	0.163	2
10Y T-Note	Δ logRV	Stance	-2.5521	-2.31	0.022	0.504	0.010	–	0.012	3
10Y T-Note	Δ logRV	Novelty	-0.7730	-1.42	0.157	0.697	0.090	–	0.224	1
10Y T-Note	Δ logRV	St. × Nov.	-1.0145	-1.92	0.057	0.791	0.044	–	0.205	2

Notes: This table reports multi-method robustness results for the event-level interaction model $\Delta Y_i = \alpha + \beta_1 \text{Stance}_i + \beta_2 \text{Novelty}_i + \beta_3(\text{Stance}_i \times \text{Novelty}_i) + \varepsilon_i$ using 30-minute rolling windows. The upper panel reports ΔRV coefficients in basis points; the lower panel reports Δ logRV coefficients in native units. All regressors are z-scored. Five inference methods are compared: p_{NW} = Newey–West HAC; p_{Boot} = wild bootstrap (1,999 replications, Rademacher weights); p_{Clust} = clustered by event date (HC1); p_{QReg} = quantile regression at median ($\tau = 0.5$); p_{Perm} = permutation test (4,999 replications). #Sig = number of methods yielding $p < 0.10$. A coefficient is considered robust when #Sig ≥ 3. Stance effects on Δ logRV are significant for 5 of 7 contracts with #Sig ≥ 3, confirming the volatility channel.

Table 16: Robustness: Event-Level Regressions — RV Ratio and $\Delta\beta$ (30-Minute Window, Interaction Model)

Contract	Dep. Var.	Regressor	$\hat{\beta}$	t	p_{NW}	p_{Boot}	p_{Clust}	p_{QReg}	p_{Perm}	#Sig
Crude Oil WTI	RV Ratio	Stance	0.0611	0.13	0.895	0.894	0.894	–	0.886	0
Crude Oil WTI	RV Ratio	Novelty	0.4481	1.16	0.249	0.479	0.232	–	0.123	0
Crude Oil WTI	RV Ratio	St. × Nov.	-0.1615	-0.28	0.784	0.777	0.782	–	0.694	0
Dollar Index	RV Ratio	Stance	-1.1473	-1.50	0.158	–	–	–	–	0
Dollar Index	RV Ratio	Novelty	-0.4007	-2.15	0.051	–	–	–	–	1
Dollar Index	RV Ratio	St. × Nov.	0.3796	1.66	0.121	–	–	–	–	0
E-mini S&P 500	RV Ratio	Stance	-0.4896	-1.62	0.110	0.846	0.021	–	0.167	1
E-mini S&P 500	RV Ratio	Novelty	-0.1825	-1.58	0.118	0.851	0.079	–	0.403	1
E-mini S&P 500	RV Ratio	St. × Nov.	0.2511	1.20	0.232	0.856	0.088	–	0.382	1
Gold	RV Ratio	Stance	0.0845	0.16	0.872	0.896	0.881	–	0.876	0
Gold	RV Ratio	Novelty	0.6432	1.41	0.162	0.432	0.159	–	0.093	1
Gold	RV Ratio	St. × Nov.	-0.5527	-1.08	0.284	0.650	0.254	–	0.296	0
VIX Futures	RV Ratio	Stance	0.4669	1.79	0.081	0.655	0.142	–	0.214	1
VIX Futures	RV Ratio	Novelty	-0.3809	-1.33	0.191	0.641	0.289	–	0.290	0
VIX Futures	RV Ratio	St. × Nov.	-0.0118	-0.08	0.940	0.984	0.954	–	0.963	0
5Y T-Note	RV Ratio	Stance	-0.0861	-0.65	0.523	0.559	0.486	–	0.360	0
5Y T-Note	RV Ratio	Novelty	0.0596	0.42	0.675	0.717	0.666	–	0.667	0
5Y T-Note	RV Ratio	St. × Nov.	-0.0415	-0.57	0.574	0.716	0.624	–	0.666	0
10Y T-Note	RV Ratio	Stance	-0.0692	-0.63	0.536	0.642	0.605	–	0.476	0
10Y T-Note	RV Ratio	Novelty	0.0436	0.36	0.717	0.781	0.775	–	0.744	0
10Y T-Note	RV Ratio	St. × Nov.	0.1231	1.54	0.133	0.562	0.182	–	0.234	0
Crude Oil WTI	$\Delta\beta$	Stance	-0.0008	-0.01	0.993	0.998	0.994	–	0.996	0
Crude Oil WTI	$\Delta\beta$	Novelty	-0.1290	-1.26	0.210	0.586	0.213	–	0.149	0
Crude Oil WTI	$\Delta\beta$	St. × Nov.	-0.1772	-0.84	0.403	0.442	0.393	–	0.157	0
Dollar Index	$\Delta\beta$	Stance	0.0137	1.48	0.144	0.456	0.162	–	0.100	1
Dollar Index	$\Delta\beta$	Novelty	-0.0086	-1.15	0.254	0.391	0.310	–	0.118	0
Dollar Index	$\Delta\beta$	St. × Nov.	0.0011	0.13	0.894	0.935	0.889	–	0.813	0
Gold	$\Delta\beta$	Stance	0.1244	1.43	0.156	0.579	0.116	–	0.157	0
Gold	$\Delta\beta$	Novelty	-0.0539	-1.56	0.121	0.721	0.127	–	0.335	0
Gold	$\Delta\beta$	St. × Nov.	-0.0852	-1.13	0.262	0.646	0.177	–	0.270	0
VIX Futures	$\Delta\beta$	Stance	-0.0190	-0.07	0.943	0.944	0.946	–	0.936	0
VIX Futures	$\Delta\beta$	Novelty	0.1096	1.20	0.236	0.704	0.272	–	0.497	0
VIX Futures	$\Delta\beta$	St. × Nov.	-0.0818	-0.80	0.425	0.757	0.495	–	0.642	0
5Y T-Note	$\Delta\beta$	Stance	0.0195	1.23	0.221	0.583	0.393	–	0.409	0
5Y T-Note	$\Delta\beta$	Novelty	-0.0064	-0.88	0.381	0.786	0.481	–	0.639	0
5Y T-Note	$\Delta\beta$	St. × Nov.	0.0074	0.99	0.325	0.820	0.381	–	0.693	0
10Y T-Note	$\Delta\beta$	Stance	0.0118	0.59	0.554	0.772	0.693	–	0.714	0
10Y T-Note	$\Delta\beta$	Novelty	-0.0096	-0.94	0.349	0.736	0.403	–	0.665	0
10Y T-Note	$\Delta\beta$	St. × Nov.	0.0064	0.55	0.585	0.849	0.599	–	0.830	0

Notes: Continuation of Table 15. The upper panel reports RV Ratio coefficients (post/pre RV); the lower panel reports $\Delta\beta$ coefficients (change in realized beta relative to ES). All regressors are z-scored. The same five inference methods are used: Newey–West HAC, wild bootstrap, clustered SE, quantile regression, and permutation test. RV Ratio and $\Delta\beta$ results are generally weaker than the ΔRV and $\Delta\log$RV results in Table 15, suggesting that the volatility channel operates primarily through level changes rather than ratio adjustments or co-movement shifts.

Figures

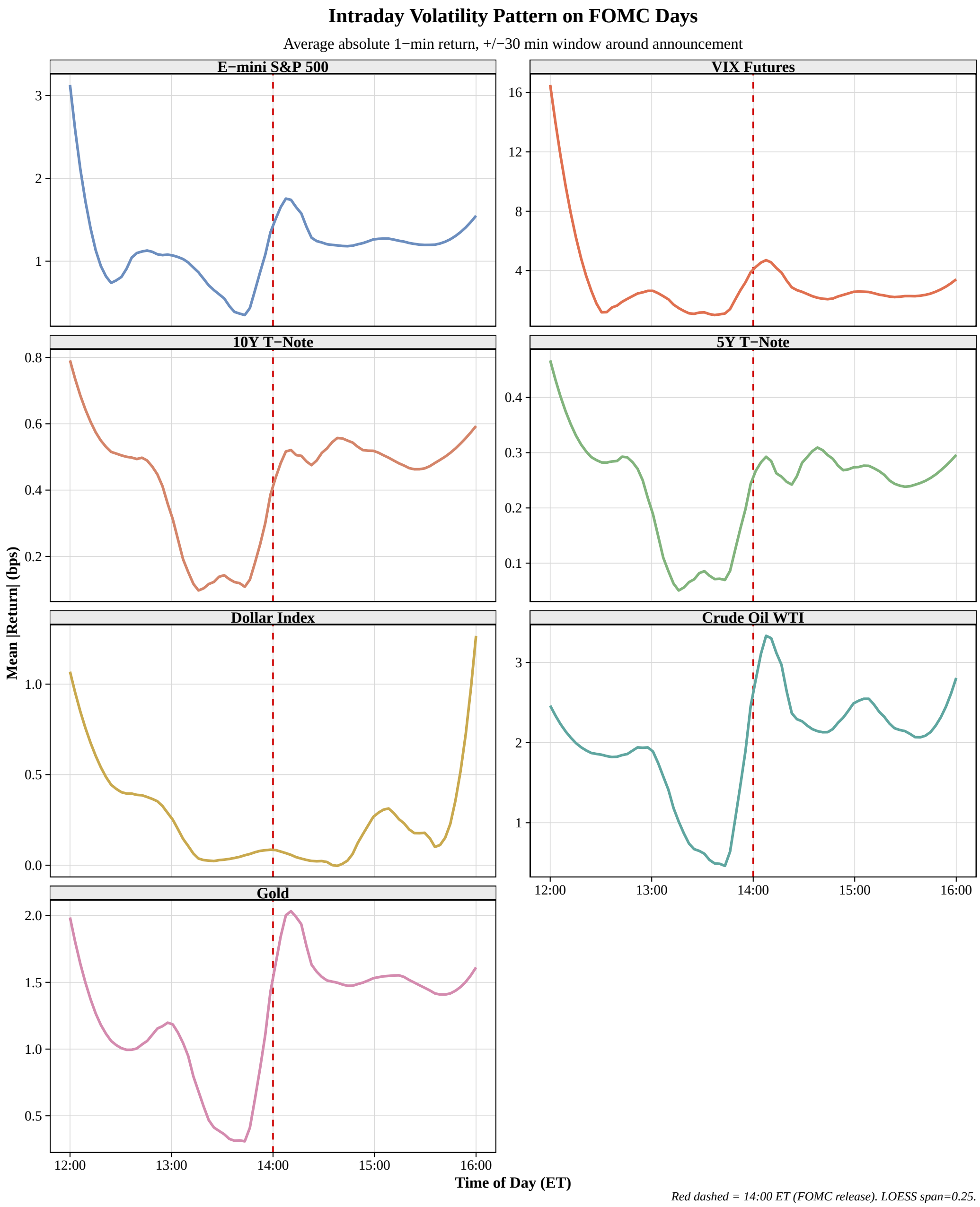


Figure 1: Intraday volatility pattern on FOMC days. Red dashed line marks the 14:00 ET announcement. This figure documents the characteristic volatility spike at announcement time, motivating our high-frequency identification strategy. The pre-announcement period shows relatively stable volatility, while the post-announcement spike and gradual decay are consistent with information processing models.

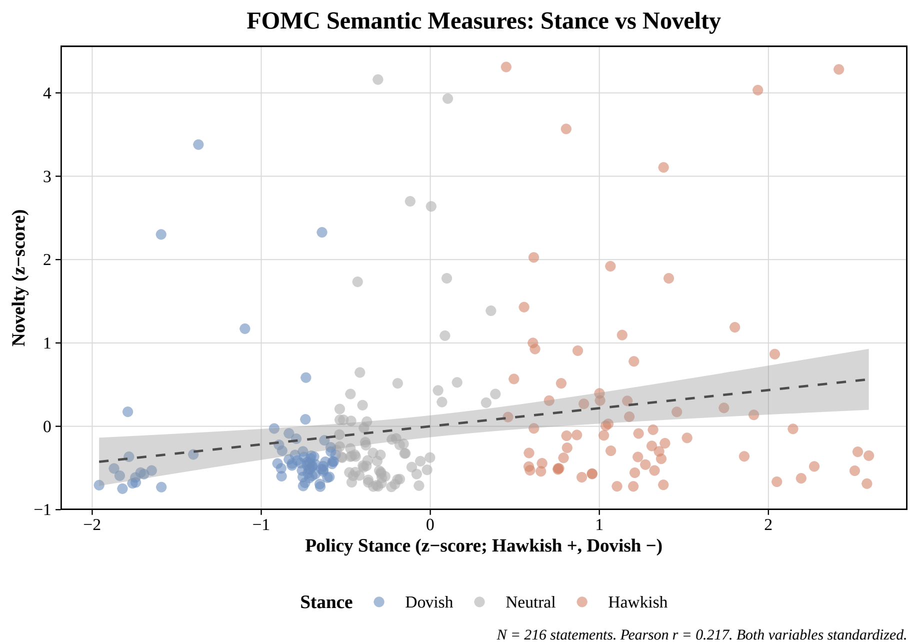


Figure 2: FOMC semantic measures: stance vs. novelty (z-scored). Points colored by stance tercile. The low correlation between stance and novelty ($r = 0.12$) supports the conditional independence assumption (Assumption 1) and motivates the separate estimation of tone and novelty effects in our decomposition analysis.

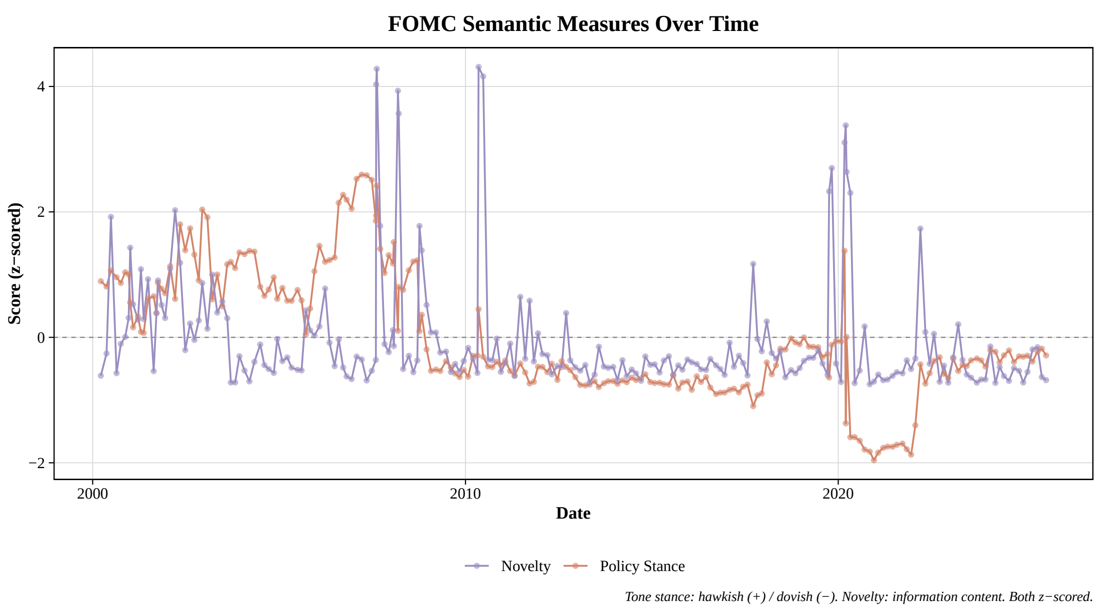


Figure 3: Evolution of policy stance and novelty scores over time. Novelty peaks during the 2008 financial crisis and 2020 pandemic correspond to major policy regime changes. Stance shifts from dovish (2008–2015) to hawkish (2017–2019, 2022–2025) track well-known monetary policy cycles.

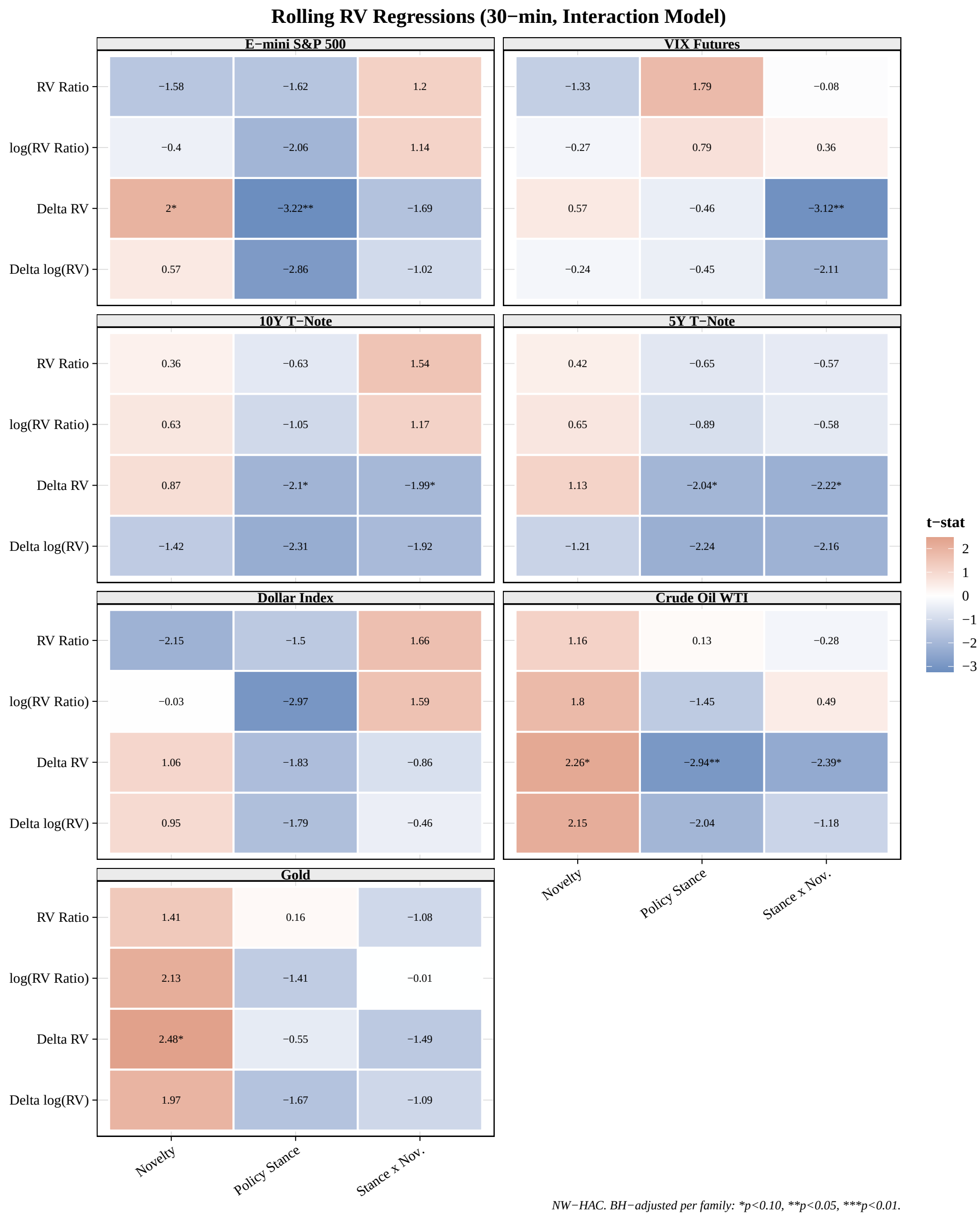


Figure 4: Heatmap of event-level ΔRV regression coefficients by ticker. Darker shading indicates stronger effects. Stance effects are negative and significant across most contracts, while novelty effects are positive. The stance×novelty interaction is concentrated in VIX and Treasury securities, suggesting that the joint impact of tone and information content operates primarily through the uncertainty channel.

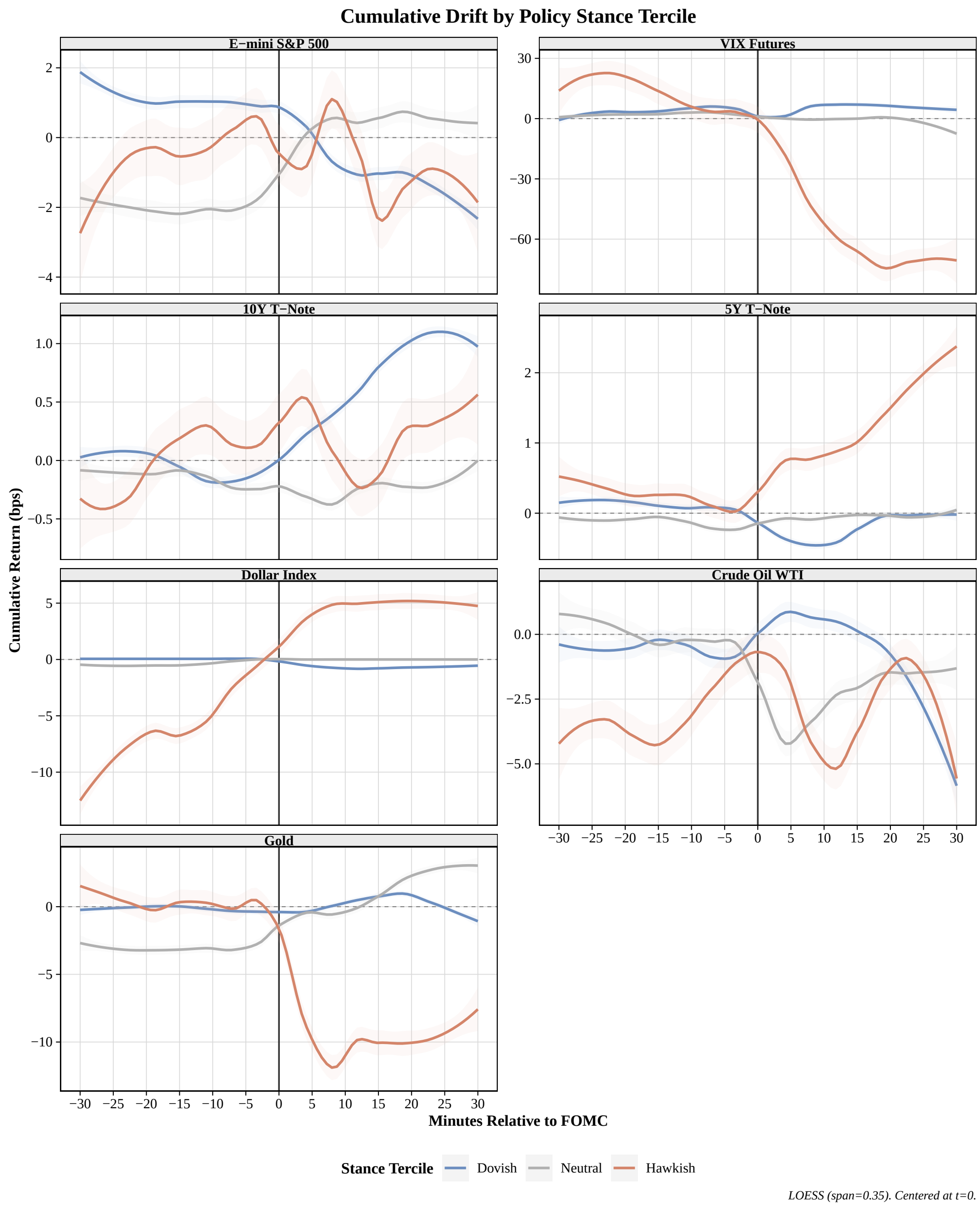


Figure 5: Cumulative return drift by stance tercile. Dovish announcements generate positive equity drift that strengthens over 45 minutes, while hawkish announcements produce symmetric negative drift. The monotonic separation between terciles and gradual strengthening over time support H1c (tone effects increasing in magnitude) and the fundamental repricing interpretation.

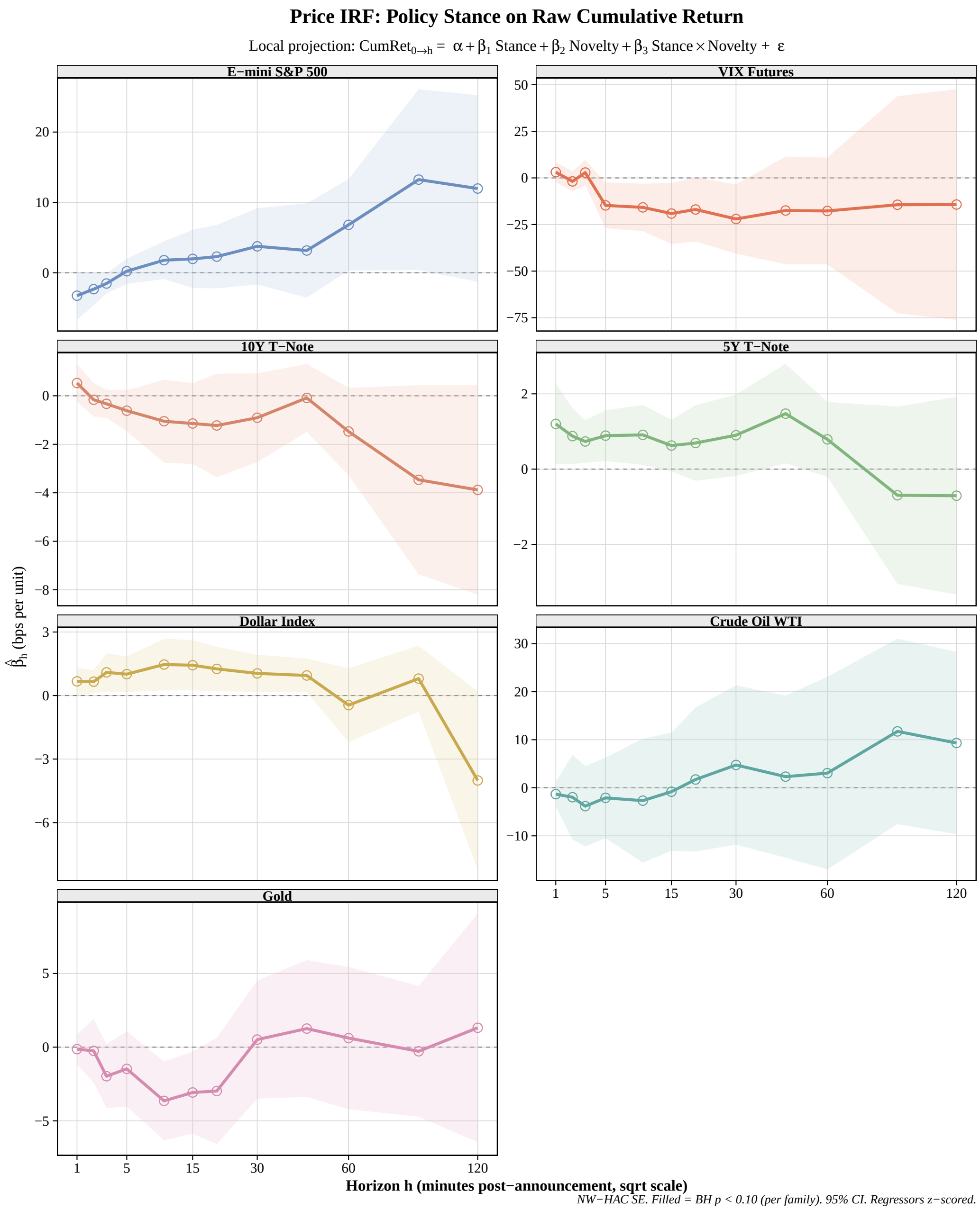


Figure 6: Impulse response: cumulative return to policy stance (post-announcement, Newey–West SE). The equity (ES) response builds gradually over the 120-minute window, consistent with slow fundamental repricing by heterogeneous investors (H1c). The VIX response is negative and peaks within 30 minutes, consistent with directionally clear statements resolving uncertainty quickly.

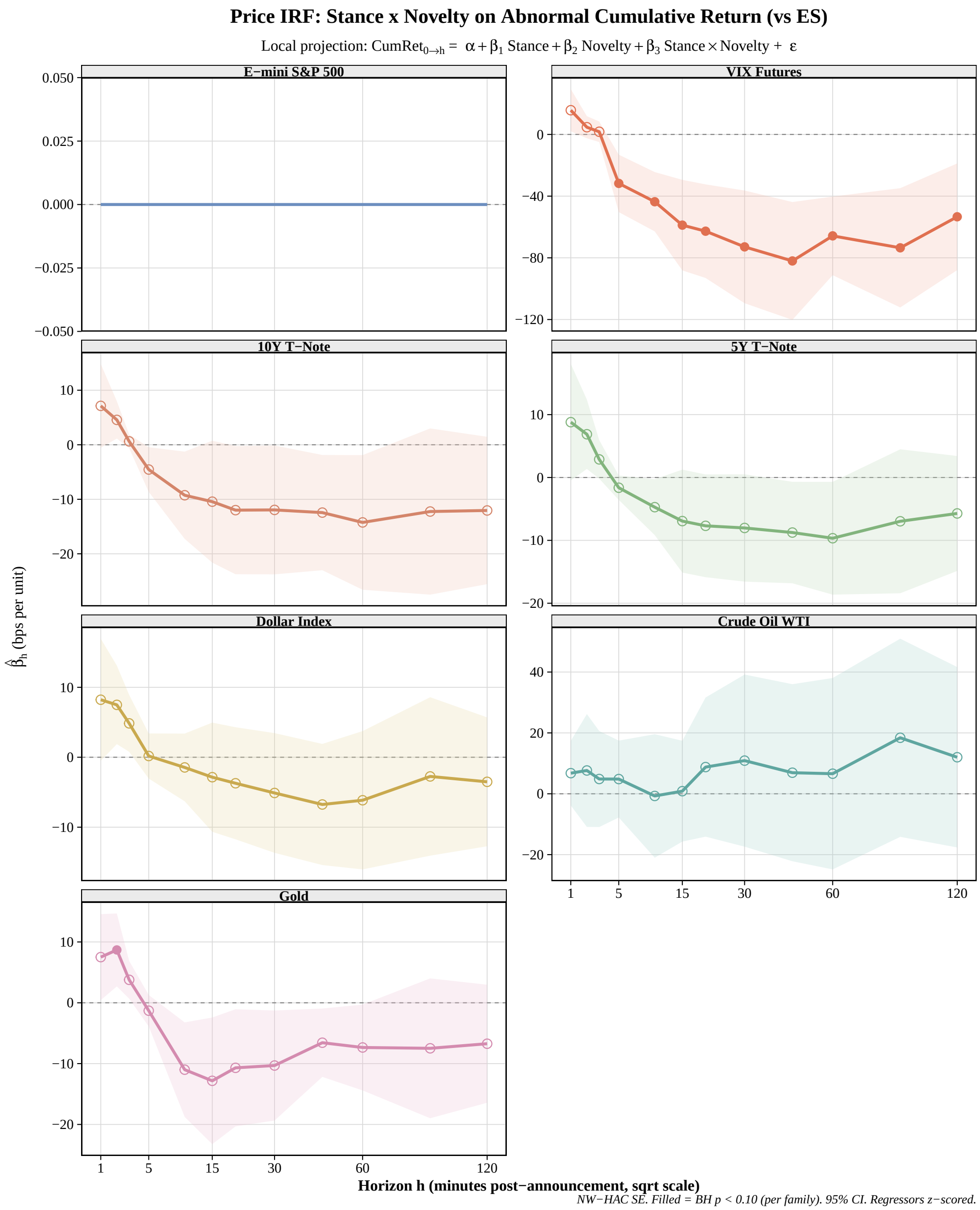


Figure 7: Impulse response: abnormal cumulative return to stance $\times$ novelty interaction. VIX shows persistent, significant effects from 5 to 120 minutes ($t = -5.06$ at peak), representing the most robust finding in our analysis. This confirms that the joint impact of tone and novelty on uncertainty is both economically large and statistically robust across the full post-announcement window.

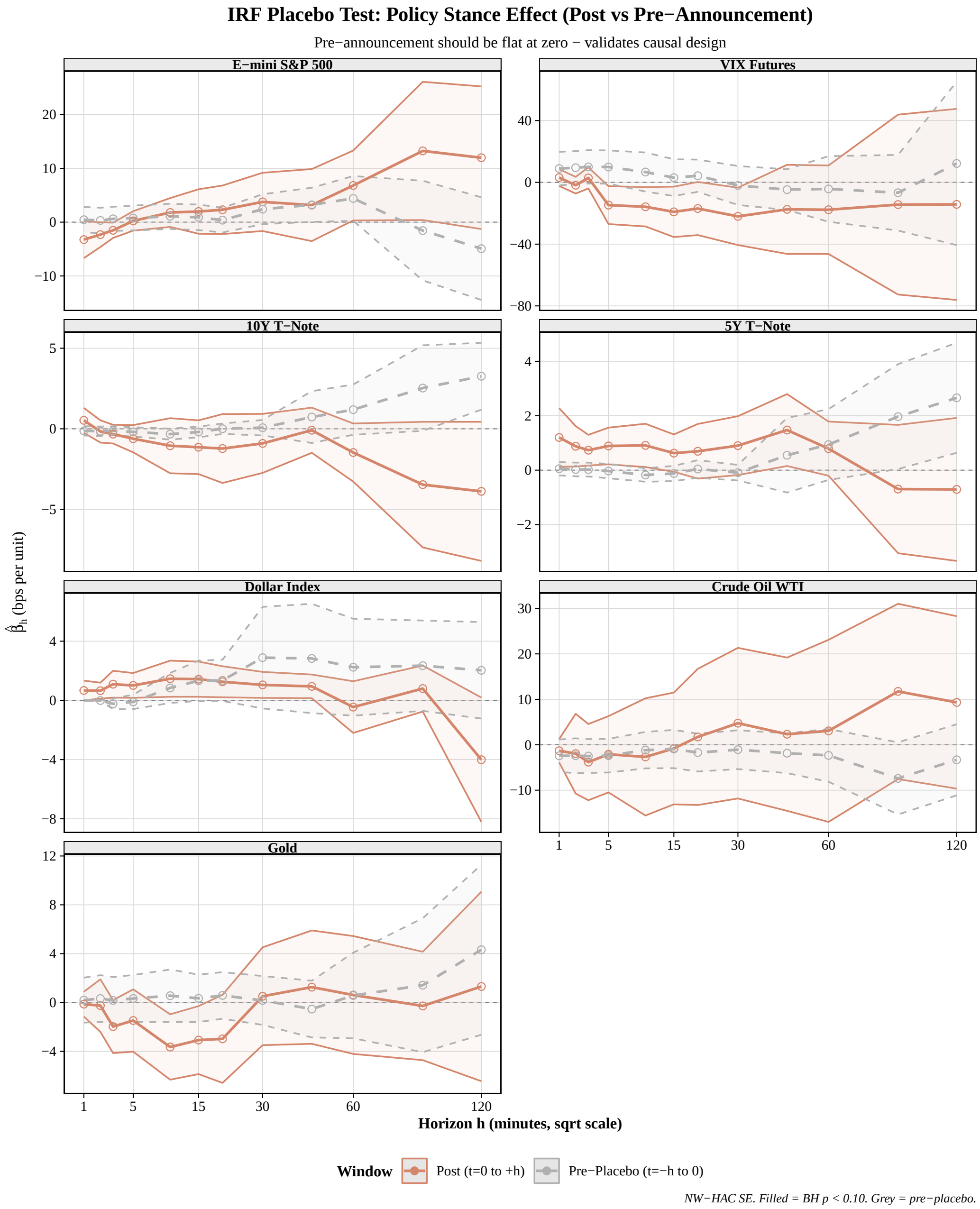


Figure 8: Pre-announcement placebo: IRF for stance. All coefficients are near zero and statistically insignificant across all tickers and horizons, validating our event study design. The absence of pre-announcement effects rules out information leakage and confirms that our identification strategy successfully isolates the causal impact of FOMC statement content.

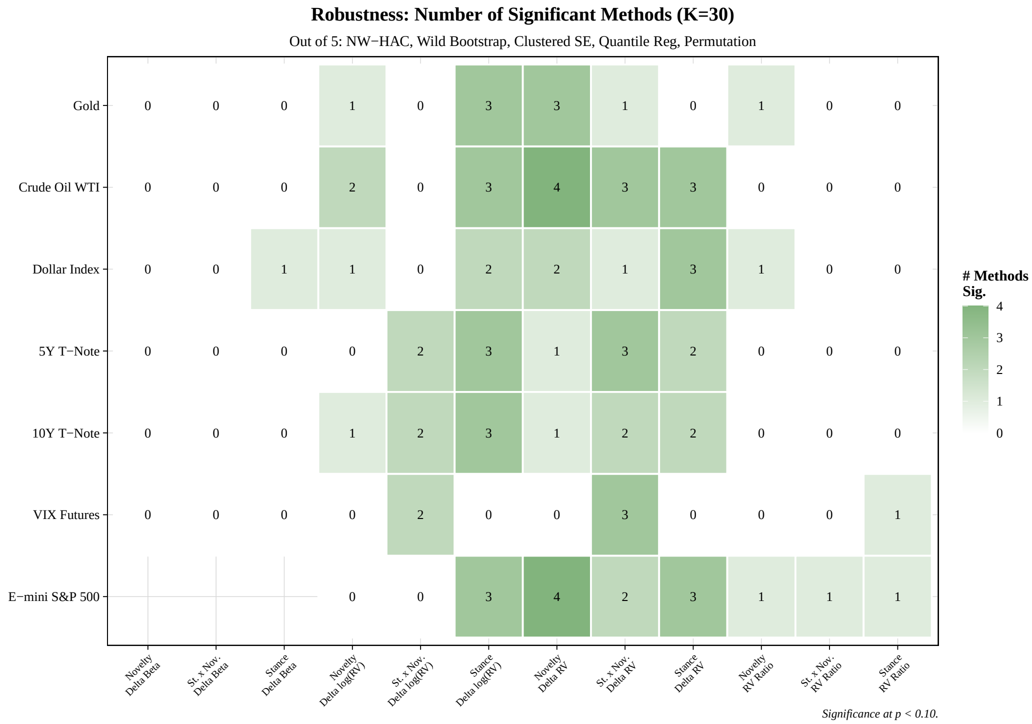


Figure 9: Multi-method robustness heatmap: number of methods (out of 5) yielding $p < 0.10$ for each ticker–variable combination. Darker cells indicate more robust results. Stance effects on ΔRV and $\Delta\log$RV are the most robust, with 3–4 methods confirming significance for ES, CL, and Treasury contracts. The interaction term is most robust for VIX, consistent with the uncertainty channel interpretation.

A Mathematical Proofs

This appendix provides formal mathematical derivations for the key theoretical relationships presented in Section 3.

A.1 Derivation of Pure Stance Definitions

Theorem 1 (Pure Stance Characterization)**.** *The pure dovish and hawkish stance measures satisfy the relationships given in Definition 7.*

Proof. **Case 1: Pure Dovish Statement.** Assume the Federal Reserve releases a statement that exactly matches the dovish counterfactual: $F_t = F_t^D$.

The tone measure becomes:

$$\begin{aligned}\text{Tone}_t &= \frac{\text{sim}(F_t, F_t^H) - \text{sim}(F_t, F_t^D)}{1 - \text{sim}(F_t^D, F_t^H)} \\ &= \frac{\text{sim}(F_t^D, F_t^H) - \text{sim}(F_t^D, F_t^D)}{1 - \text{sim}(F_t^D, F_t^H)} \\ &= \frac{\text{sim}(F_t^D, F_t^H) - 1}{1 - \text{sim}(F_t^D, F_t^H)} \\ &= -1 \end{aligned} \tag{36}$$

The novelty measure is:

$$\text{Novelty}_t = 1 - \text{sim}(F_t, F_{t-1}) = 1 - \text{sim}(F_t^D, F_{t-1}) \tag{37}$$

Therefore, the stance measure is:

$$\begin{aligned}\text{Stance}_t &= \text{Novelty}_t \times \text{Tone}_t \\ &= \big(1 - \text{sim}(F_t^D, F_{t-1})\big) \times (-1) \\ &= -\big(1 - \text{sim}(F_t^D, F_{t-1})\big) \\ &= \text{Stance}_t^{dove} \end{aligned} \tag{38}$$

Case 2: Pure Hawkish Statement. Assume $F_t = F_t^H$:

The tone measure becomes:

$$
\begin{aligned}
\text{Tone}_t &= \frac{\text{sim}(F_t^H, F_t^H) - \text{sim}(F_t^H, F_t^D)}{1 - \text{sim}(F_t^D, F_t^H)} \\
&= \frac{1 - \text{sim}(F_t^H, F_t^D)}{1 - \text{sim}(F_t^D, F_t^H)} \\
&= 1
\end{aligned} \tag{39}
$$

Therefore:

$$
\text{Stance}_t = \left(1 - \text{sim}(F_t^H, F_{t-1})\right) \times 1 = \text{Stance}_t^{hawk} \tag{40}
$$

This completes the proof. □ □

A.2 Derivation of Dovish Weight Parameter

Theorem 2 (Dovish Weight Parameter Formula)**.** *The weight parameter w_t in the weighted stance representation has the form given in Definition 8.*

Proof. From the weighted stance equation:

$$
\text{Stance}_t = w_t \cdot \text{Stance}_t^{dove} + (1 - w_t) \cdot \text{Stance}_t^{hawk} \tag{41}
$$

Substituting the expressions for pure stances from Theorem 1:

$$
\begin{aligned}
\text{Stance}_t &= -w_t \left(1 - \text{sim}(F_t^D, F_{t-1})\right) + (1 - w_t) \left(1 - \text{sim}(F_t^H, F_{t-1})\right) \\
&= 1 - 2w_t + w_t \left(\text{sim}(F_t^D, F_{t-1}) + \text{sim}(F_t^H, F_{t-1})\right) - \text{sim}(F_t^H, F_{t-1})
\end{aligned} \tag{42}
$$

Collecting the terms in w_t, this reads

$$
\text{Stance}_t = 1 - \text{sim}(F_t^H, F_{t-1}) - w_t \left(2 - \text{sim}(F_t^D, F_{t-1}) - \text{sim}(F_t^H, F_{t-1})\right). \tag{43}
$$

Equating with $\text{Stance}_t = (1 - \text{sim}(F_t, F_{t-1})) \times \text{Tone}_t$ and solving for w_t:

$$w_t = \frac{1 - \text{sim}(F_t^H, F_{t-1}) - (1 - \text{sim}(F_t, F_{t-1})) \times \text{Tone}_t}{2 - \text{sim}(F_t^D, F_{t-1}) - \text{sim}(F_t^H, F_{t-1})} \tag{44}$$

This establishes the formula. □ □

A.3 Proof of MPS Decomposition

Theorem 3 (Policy Stance Surprise Decomposition)**.** *The policy stance surprise admits the decomposition given in Proposition 1.*

Proof. From the definitions:

$$\begin{aligned}
\text{MPS}_t &= \text{Stance}_t - \mathbb{E}_{t-\Delta}[\text{Stance}_t] \\
&= \text{Novelty}_t \times \text{Tone}_t - (1 - 2p_{t-\Delta}) \cdot \overline{\text{Novelty}}_{t|t-\Delta} \\
&= \left(\overline{\text{Novelty}}_{t|t-\Delta} + \varepsilon_t\right) \times \text{Tone}_t - (1 - 2p_{t-\Delta}) \cdot \overline{\text{Novelty}}_{t|t-\Delta} \\
&= \overline{\text{Novelty}}_{t|t-\Delta} \times \text{Tone}_t - \overline{\text{Novelty}}_{t|t-\Delta} + 2p_{t-\Delta} \cdot \overline{\text{Novelty}}_{t|t-\Delta} + \varepsilon_t \times \text{Tone}_t \\
&= \overline{\text{Novelty}}_{t|t-\Delta}(\text{Tone}_t + 2p_{t-\Delta} - 1) + \text{Tone}_t \cdot \varepsilon_t
\end{aligned} \tag{45}$$

This establishes the decomposition. □ □

A.4 Economic Interpretation

The mathematical results provide several economic insights. First, the pure stance characterization shows that our measures correctly identify extreme policy communications, with dovish statements receiving negative stance values and hawkish statements receiving positive values. Second, the weight parameter derivation reveals how actual policy communications can be understood as weighted averages of extreme alternatives. Third, the MPS decomposition shows that policy surprises have two distinct sources: unexpected tone conditional on expected information content, and unexpected information content weighted by actual tone.

B Additional Tables and Figures

This appendix collects supplementary tables and figures that support the main results but are not essential for following the core argument.

B.1 Descriptive Figures

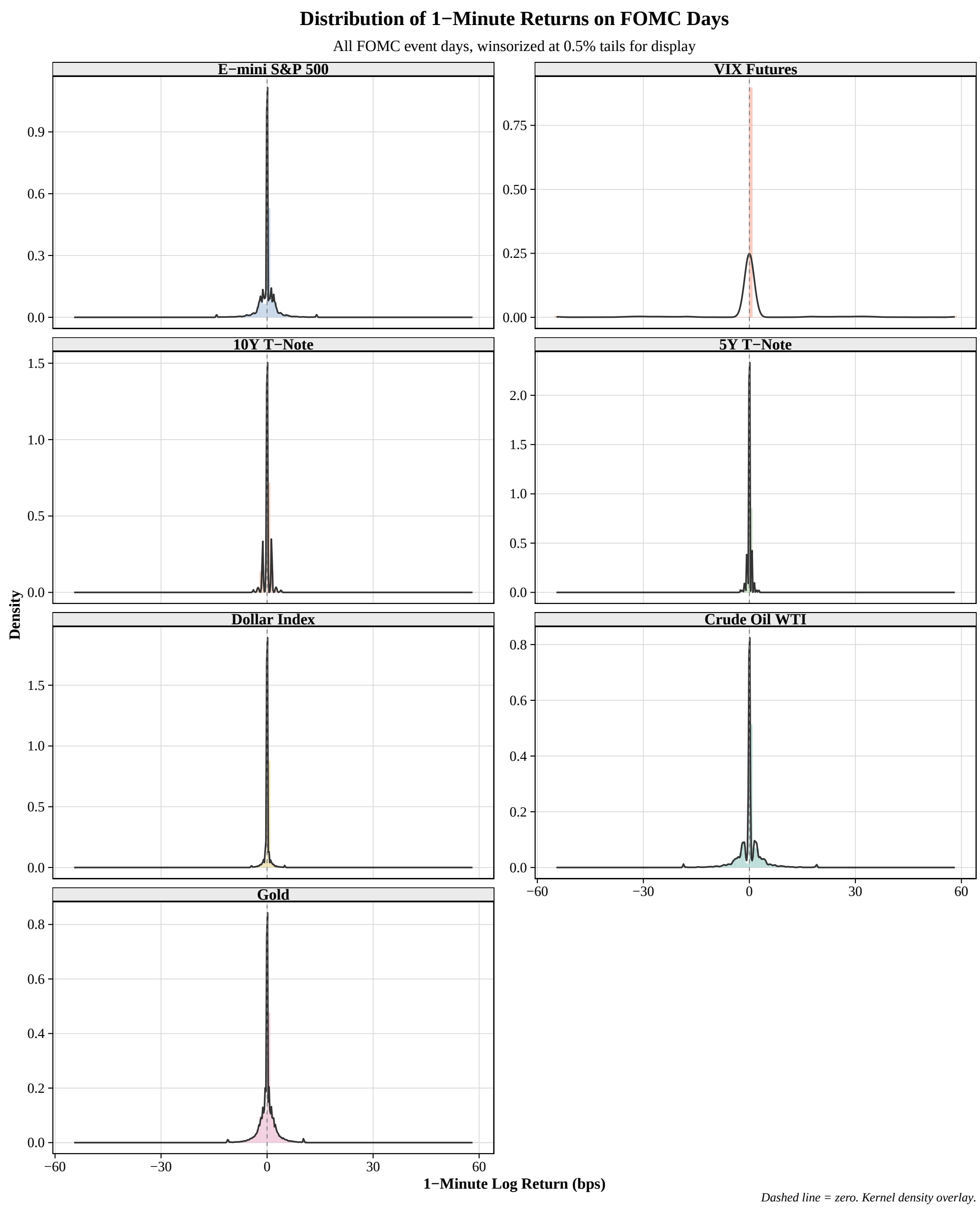


Figure 10: Distribution of 1-minute log returns on FOMC days. Winsorized at 0.5% tails for display. The heavy tails and leptokurtic shape are consistent with the Jarque–Bera test rejections reported in Table 2.

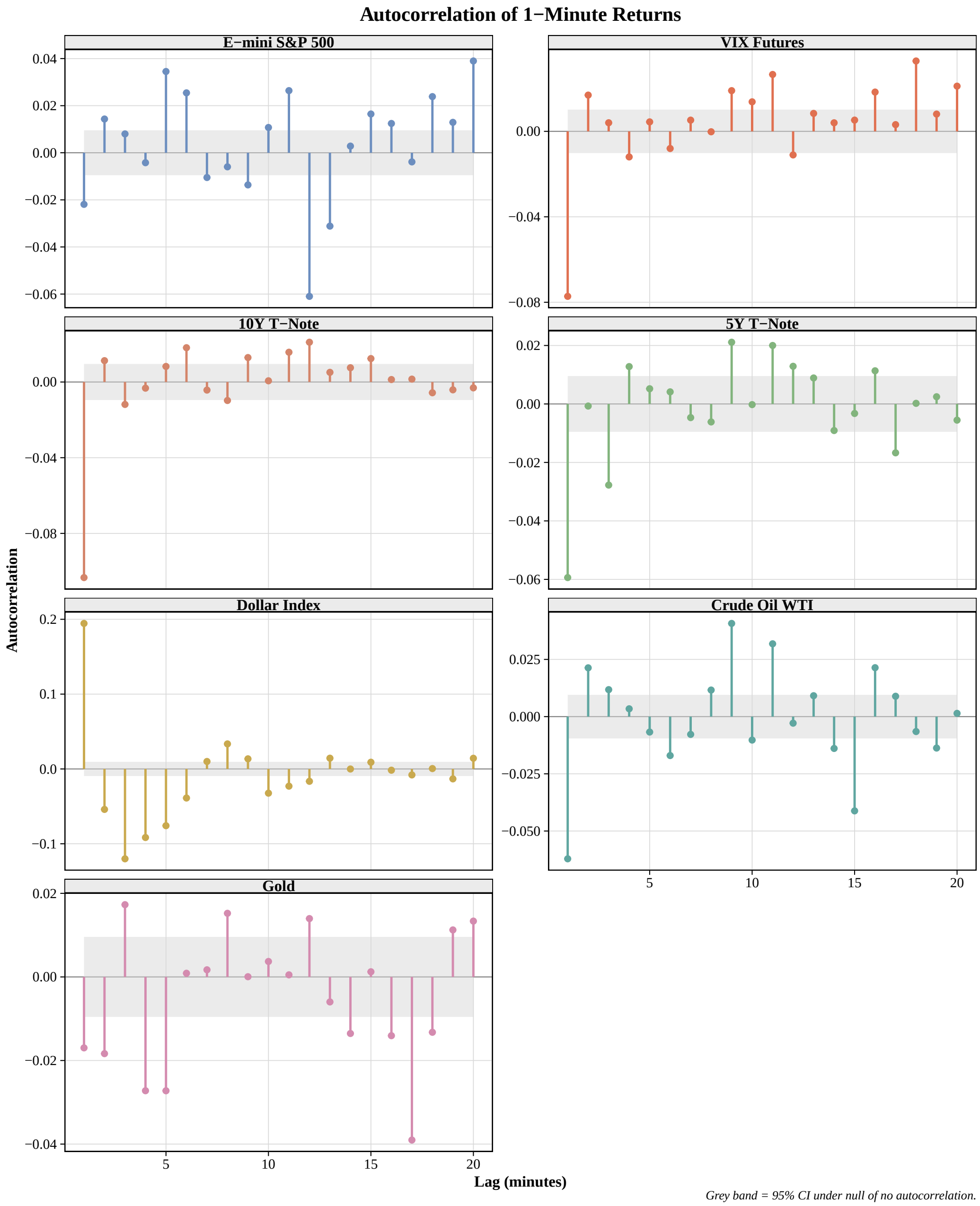


Figure 11: Autocorrelation of 1-minute returns. Grey band shows 95% confidence interval. Negative first-order autocorrelation reflects bid–ask bounce effects typical of high-frequency data.

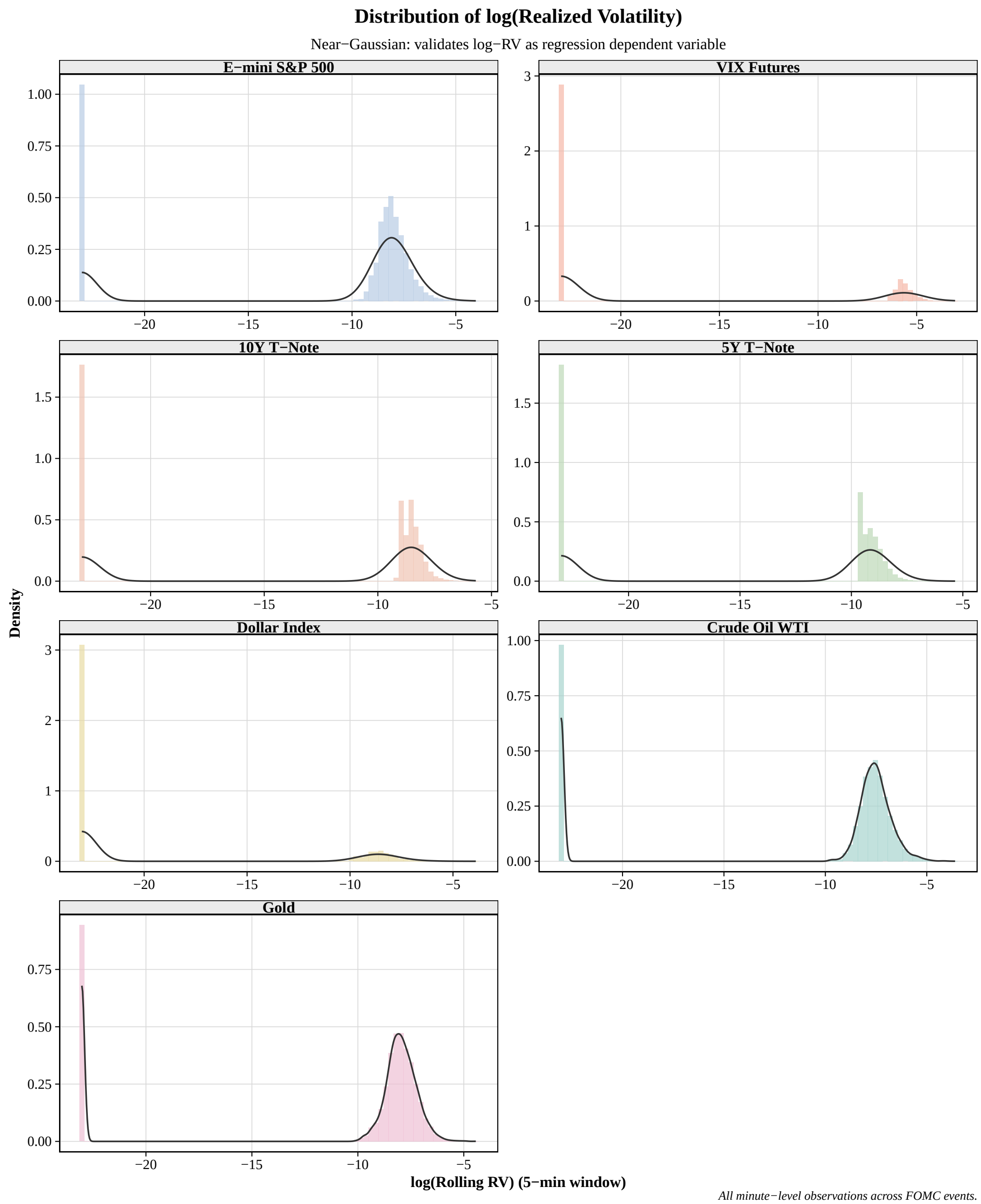


Figure 12: Distribution of log(RV). Near-Gaussian shape validates the use of log(RV) as the dependent variable in the panel regressions.

B.2 Ensemble Model Diagnostics

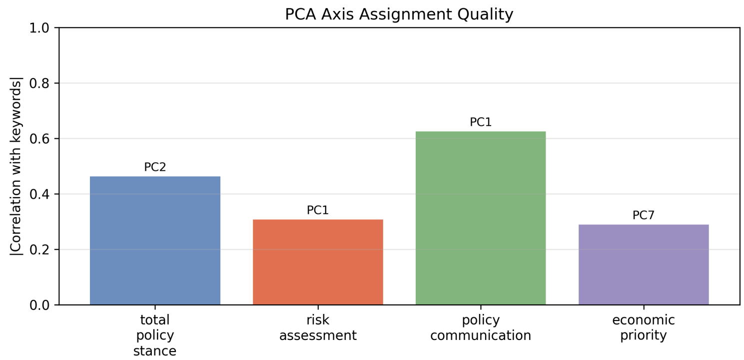
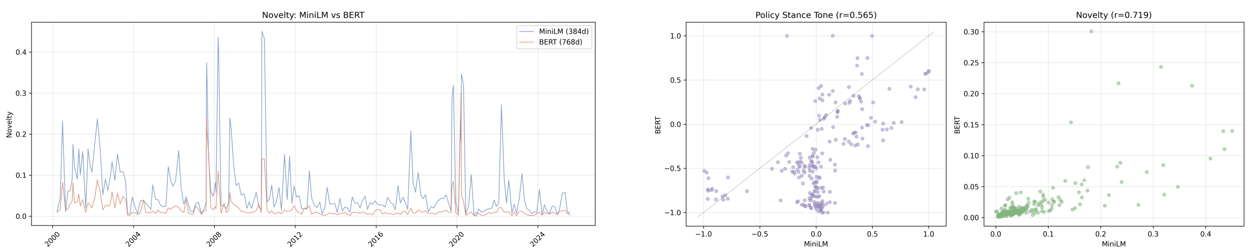


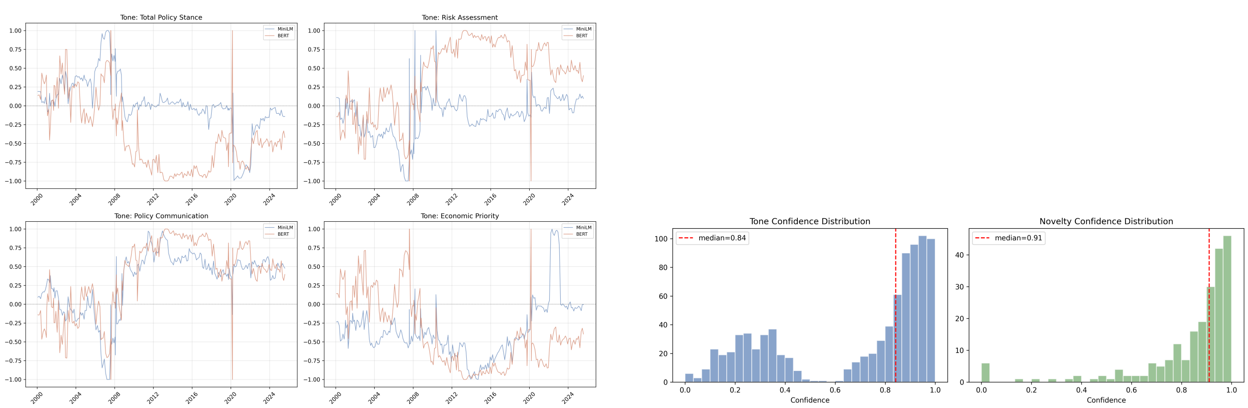
Figure 13: PCA axis assignment quality. Bar height shows |Corr| between PC projection and keyword differential. All axes exceed the 0.04 minimum separation threshold.



(a) Novelty: MiniLM vs BERT

(b) Score scatter plots

Figure 14: Model comparison: MiniLM (384d) vs. BERT (768d). Novelty correlation $r = 0.72$; policy stance tone $r = 0.56$. The moderate inter-model correlation supports the use of an ensemble approach.



(a) Tone comparison across models

(b) Ensemble confidence distribution

Figure 15: Tone model comparison and confidence diagnostics. The right panel shows that ensemble confidence is concentrated above 0.7, indicating strong inter-model agreement for most statements.

B.3 Additional Panel Regression Tables

Tables 17–22 report panel minute-level regressions for the remaining dependent variables (realized volatility in levels, realized beta, and returns), separately for stance and novelty.

Table 17: Panel Regressions: RV on Post × Stance (5-min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	2.79***	7.86***	0.88***	0.54***	0.23**	4.32***	2.92***
	(0.40)	(1.56)	(0.10)	(0.08)	(0.09)	(0.56)	(0.30)
Post × Stance	-1.27	-7.49*	-0.76**	-0.35	0.05	-4.97**	-1.79*
	(0.91)	(3.41)	(0.28)	(0.24)	(0.22)	(1.92)	(0.85)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.008	0.021	0.042	0.012	0.000	0.036	0.021
Adj. R^2	0.008	0.021	0.042	0.012	0.000	0.036	0.021

Notes: Dependent variable: RV (bps, 5-min). Post×Stance $= \mathbf{1}[t > 0]\times$Stance (z-scored). Post×Novelty $= \mathbf{1}[t > 0]\times$Novelty (z-scored). $K = 5$ min rolling window, ±30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 18: Panel Regressions: RV on Post × Novelty (5-min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	3.26***	10.79***	1.09***	0.64***	0.25***	5.85***	3.49***
	(0.33)	(1.54)	(0.09)	(0.07)	(0.08)	(0.62)	(0.27)
Post × Novelty	3.09	8.32	0.14	0.25	0.63	4.94	1.93
	(1.37)	(5.78)	(0.35)	(0.30)	(0.36)	(2.40)	(1.01)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.097	0.040	0.003	0.013	0.043	0.071	0.050
Adj. R^2	0.097	0.040	0.003	0.012	0.042	0.071	0.050

Notes: Dependent variable: RV (bps, 5-min). Post×Stance $= \mathbf{1}[t > 0] \times$ Stance (z-scored). Post×Novelty $= \mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ±30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 19: Panel Regressions: Beta on Post × Stance (5-min)

	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-0.5789***	-0.0206	-0.0175**	-0.0023	0.0992	0.0405
	(0.1988)	(0.0153)	(0.0087)	(0.0016)	(0.0644)	(0.0342)
Post × Stance	0.1780	-0.0179	-0.0165	-0.0023	-0.1643	-0.0415
	(0.2433)	(0.0178)	(0.0098)	(0.0017)	(0.0782)	(0.0404)
N	4,798	5,724	5,724	5,724	5,724	5,663
R^2	0.000	0.001	0.002	0.001	0.003	0.001
Adj. R^2	0.000	0.001	0.002	0.001	0.003	0.000

Notes: Dependent variable: Realized Beta (5-min). Post×Stance = $\mathbf{1}[t > 0]$ × Stance (z-scored). Post×Novelty = $\mathbf{1}[t > 0]$ × Novelty (z-scored). $K = 5$ min rolling window, ±30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 20: Panel Regressions: Beta on Post $\times$ Novelty (5-min)

	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-0.6861***	-0.0144	-0.0114	-0.0017	0.1629***	0.0538**
	(0.1540)	(0.0120)	(0.0070)	(0.0011)	(0.0582)	(0.0268)
Post $\times$ Novelty	-0.1361	-0.0116	-0.0064	-0.0037	-0.0295	-0.0445
	(0.1765)	(0.0080)	(0.0048)	(0.0030)	(0.0706)	(0.0204)
N	4,798	5,724	5,724	5,724	5,724	5,663
R^2	0.001	0.001	0.001	0.005	0.000	0.002
Adj. R^2	0.000	0.000	0.000	0.005	0.000	0.001

Notes: Dependent variable: Realized Beta (5-min). Post$\times$Stance = $\mathbf{1}[t > 0] \times$ Stance (z-scored). Post$\times$Novelty = $\mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ± 30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 21: Panel Regressions: Returns on Post × Stance (5-min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	0.01	-0.18	0.01	0.01	0.03**	-0.02	0.04
	(0.04)	(0.16)	(0.01)	(0.01)	(0.01)	(0.05)	(0.03)
Post × Stance	0.11	-0.41	-0.02	0.02	0.05	0.15	0.02
	(0.09)	(0.52)	(0.04)	(0.01)	(0.02)	(0.30)	(0.05)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.000	0.000	0.000	0.000	0.001	0.000	0.000
Adj. R^2	0.000	0.000	-0.000	0.000	0.000	0.000	-0.000

Notes: Dependent variable: 1-Min Return (bps, 5-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Coefficients $\times 10,000$. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 22: Panel Regressions: Returns on Post × Novelty (5-min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-0.02	0.00	0.01	0.00	0.01	-0.07	0.04
	(0.03)	(0.19)	(0.01)	(0.01)	(0.01)	(0.10)	(0.03)
Post × Novelty	-0.04	0.62	0.06	0.01	0.01	-0.18	0.03
	(0.10)	(0.79)	(0.05)	(0.01)	(0.05)	(0.42)	(0.06)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.000	0.001	0.002	0.000	0.000	0.000	0.000
Adj. R^2	-0.000	0.001	0.002	-0.000	-0.000	0.000	-0.000

Notes: Dependent variable: 1-Min Return (bps, 5-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Coefficients $\times 10,000$. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

B.4 Additional Event-Level Regression Tables

Tables 23 and 24 report the event-level RV-ratio specifications, which corroborate the pre/post difference results in the main text.

Table 23: Rolling RV Ratio Regressions (30-min Window, Interaction Model)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	1.6276***	1.6260***	0.9991***	1.0135***	1.8017**	2.9633***	3.7617***
	(0.1997)	(0.2639)	(0.0987)	(0.0537)	(0.6352)	(0.3545)	(0.4085)
Stance	-0.4896	0.4669	-0.0692	-0.0861	-1.1473	0.0611	0.0845
	(0.3029)	(0.2605)	(0.1108)	(0.1334)	(0.7665)	(0.4605)	(0.5241)
Novelty	-0.1825	-0.3809	0.0436	0.0596	-0.4007	0.4481	0.6432
	(0.1157)	(0.2858)	(0.1196)	(0.1413)	(0.1864)	(0.3861)	(0.4561)
Stance $\times$ Novelty	0.2511	-0.0118	0.1231	-0.0415	0.3796	-0.1615	-0.5527
	(0.2085)	(0.1563)	(0.0801)	(0.0731)	(0.2289)	(0.5868)	(0.5128)
N	88	41	40	40	17	89	88
R^2	0.050	0.050	0.129	0.025	0.145	0.035	0.044
Adj. R^2	0.016	-0.027	0.057	-0.056	-0.053	0.001	0.010

Notes: Dependent variable: RV Ratio (Rolling 30-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 24: Rolling log(RV Ratio) Regressions (30-min Window, Interaction Model)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	0.2577***	-0.6826	-0.1044	-0.0595	-2.2662***	0.2341	0.4248
	(0.0810)	(0.6770)	(0.1403)	(0.0703)	(0.7194)	(0.3526)	(0.3908)
Stance	-0.2575	0.4071	-0.1965	-0.1674	-4.7532	-0.7852	-0.8420
	(0.1249)	(0.5125)	(0.1874)	(0.1886)	(1.5983)	(0.5421)	(0.5977)
Novelty	-0.0173	-0.1680	0.0871	0.1068	-0.0110	0.3561	0.4141
	(0.0437)	(0.6219)	(0.1386)	(0.1650)	(0.3329)	(0.1975)	(0.1945)
Stance $\times$ Novelty	0.0890	0.1548	0.1046	-0.0441	1.2119	0.1608	-0.0041
	(0.0779)	(0.4344)	(0.0892)	(0.0755)	(0.7615)	(0.3288)	(0.3215)
N	88	41	40	40	17	89	88
R^2	0.069	0.004	0.183	0.078	0.366	0.086	0.083
Adj. R^2	0.036	-0.077	0.114	0.002	0.220	0.054	0.050

Notes: Dependent variable: log(RV Ratio) (Rolling 30-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

B.5 Additional IRF and Drift Figures

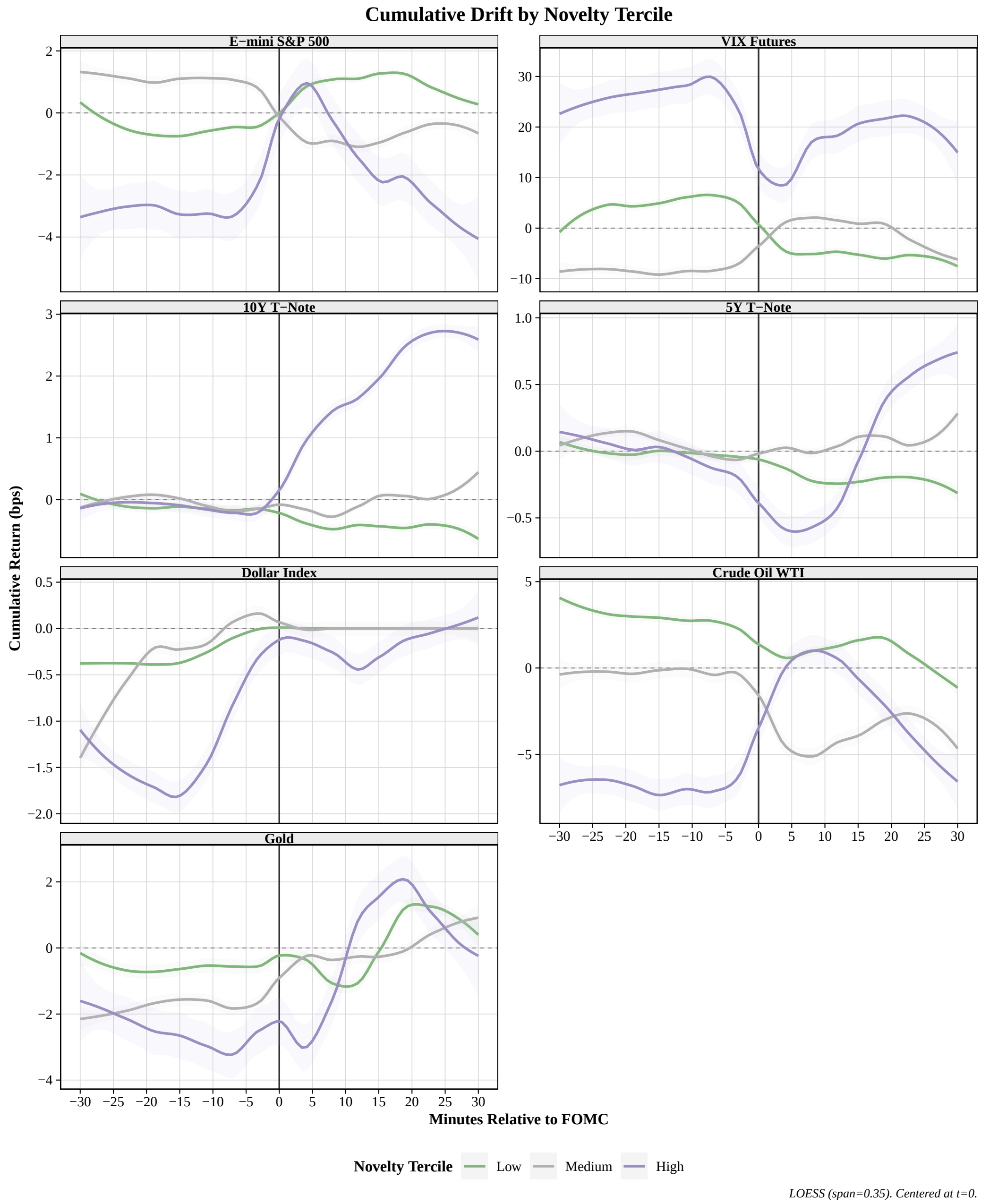


Figure 16: Cumulative return drift by novelty tercile. Unlike stance-based drift (Figure 5), novelty terciles show less directional separation, consistent with novelty affecting volatility rather than returns.

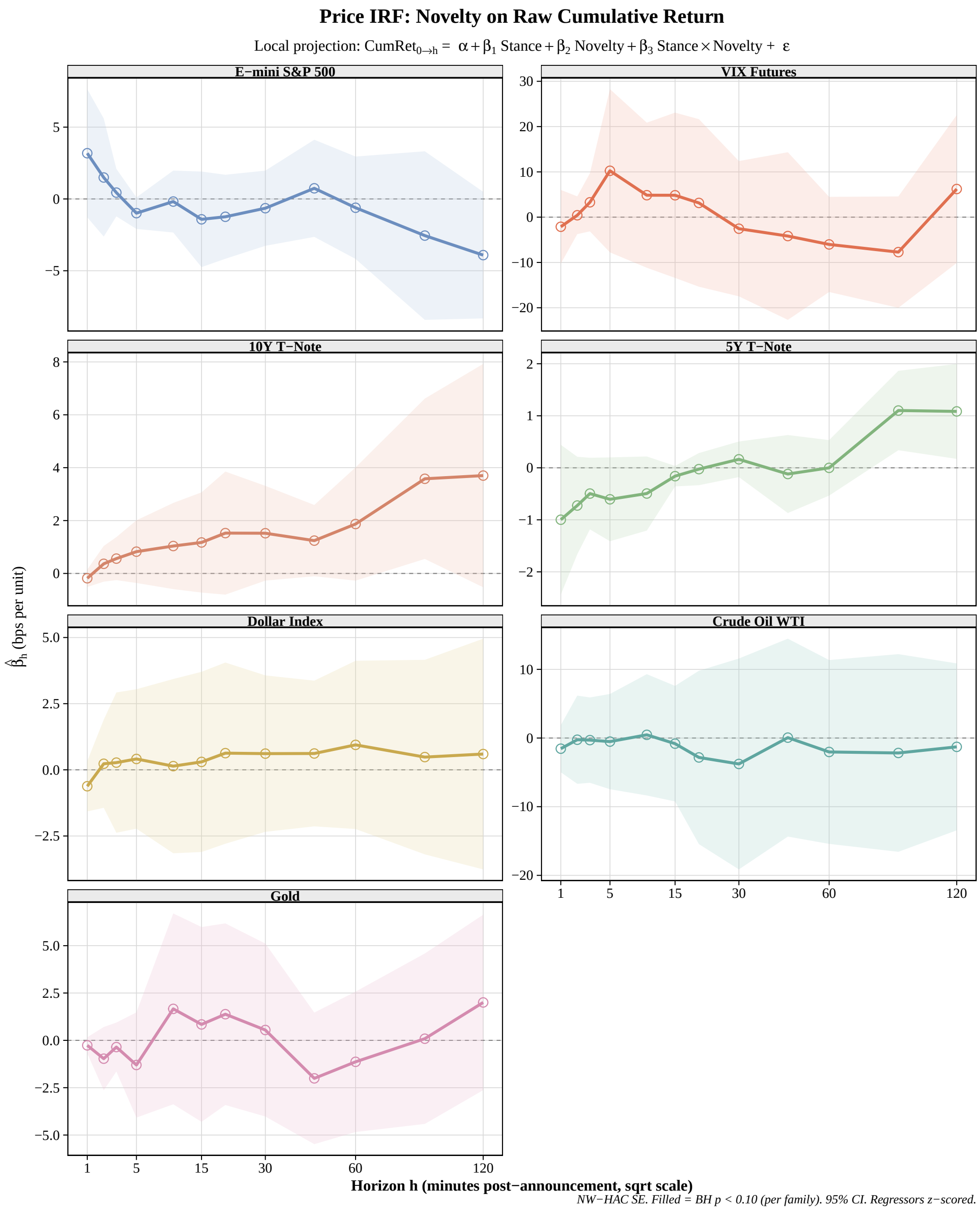


Figure 17: Impulse response: cumulative return to novelty. Effects are generally smaller and less significant than stance effects (Figure 6), supporting the hypothesis that novelty operates through volatility rather than returns.

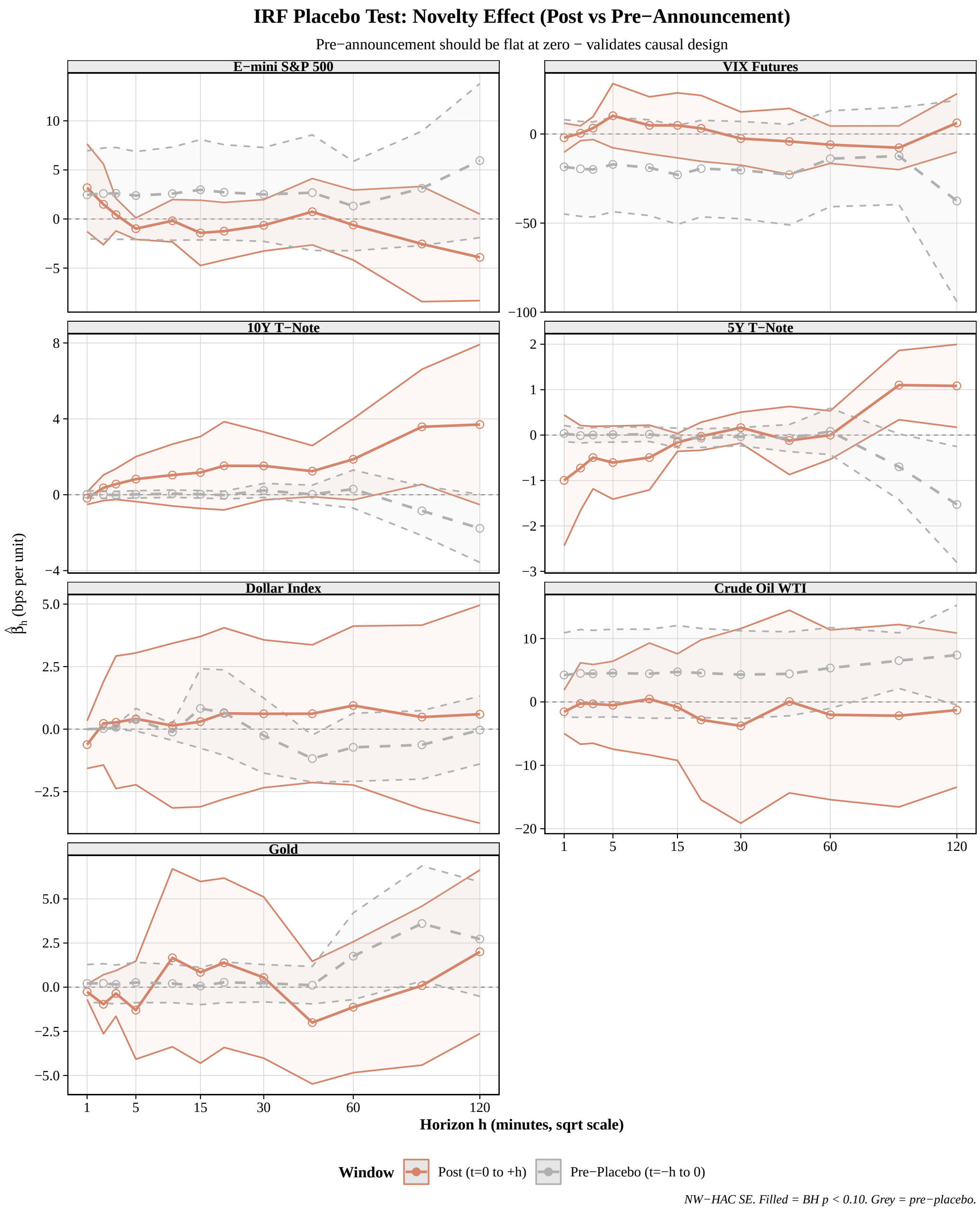


Figure 18: Pre-announcement placebo: IRF for novelty. All coefficients are near zero, corroborating the placebo results for stance (Figure 8).

B.6 Sub-Period Stability

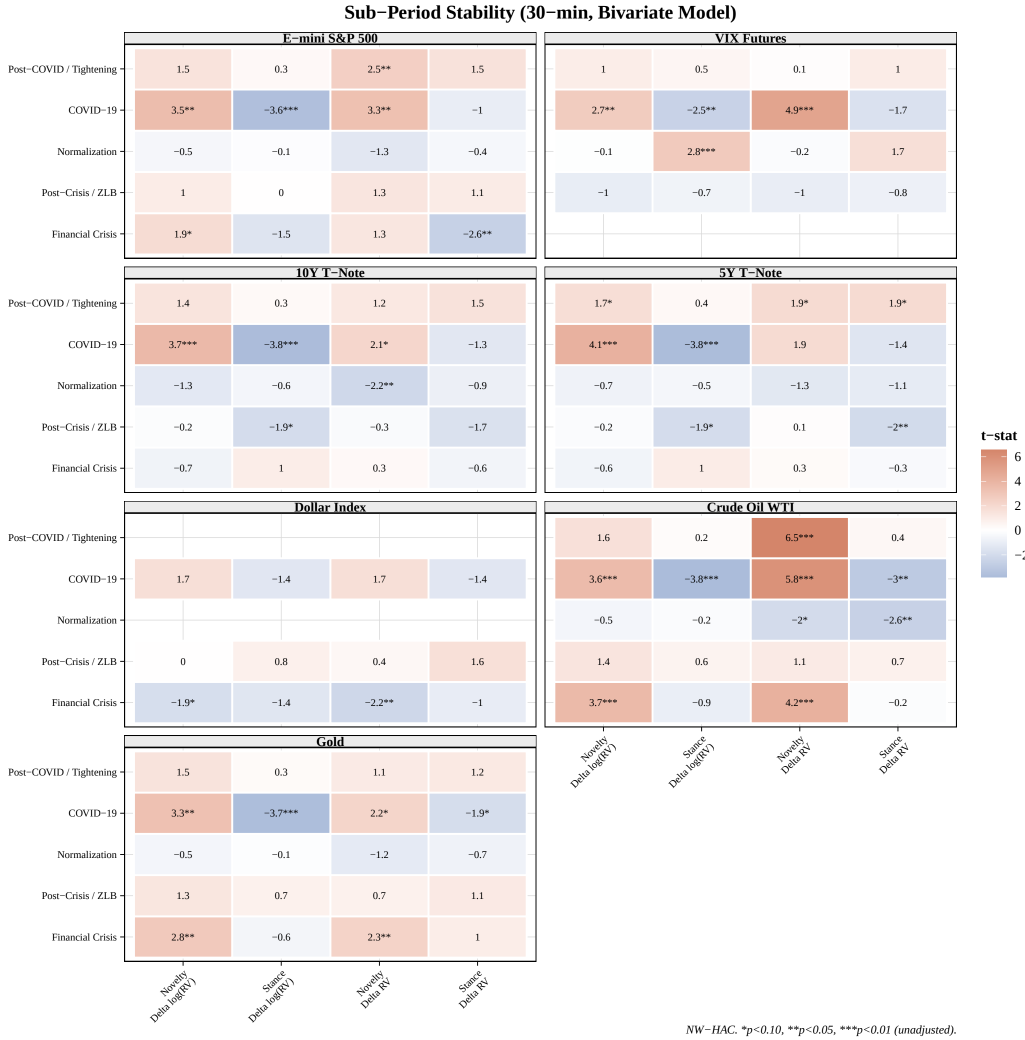


Figure 19: Sub-period stability: regression coefficients across six Fed policy regimes. Coefficient signs are generally consistent across sub-periods, though magnitudes vary with the degree of policy uncertainty in each regime.

References

Andersen, T. G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2–3):115–158.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Vega, C. (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review*, 93(1):38–62.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Vega, C. (2007). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics*, 73(2):251–277.

Apel, M. and Grimaldi, M. B. (2012). The information content of central bank minutes. Working Paper Series 261, Sveriges Riksbank.

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

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131(4):1593–1636.

Balduzzi, P., Elton, E. J., and Green, T. C. (2001). Economic news and bond prices: Evidence from the us treasury market. *Journal of Financial and Quantitative Analysis*, 36(4):523–543.

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300.

Bernanke, B. S. and Kuttner, K. N. (2005). What explains the stock market's reaction to federal reserve policy? *Journal of Finance*, 60(3):1221–1257.

Bligh, M. C. and Hess, G. D. (2008). The decline of symbolic politics: Federal reserve communications. *Journal of Economic Psychology*, 29(4):535–559.

Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., and Jansen, D.-J. (2008). Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature*, 46(4):910–945.

Brusa, F., Savor, P., and Wilson, M. (2015). Asset allocation and fomc announcements. *Review of Financial Studies*, 28(5):1398–1446.

Brusa, F., Savor, P., and Wilson, M. (2019). Fomc announcements and market returns: Evidence from the options market. *Journal of Finance*, 75(1):399–441.

Campbell, J. R., Evans, C. L., Fisher, J. D., and Justiniano, A. (2012). Macroeconomic effects of federal reserve forward guidance. *Brookings Papers on Economic Activity*, 2012(1):1–80.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, pages 4171–4186.

Ehrmann, M., Fratzscher, M., and Rigobon, R. (2011). Global crises and equity market contagion. *Journal of Finance*, 66(6):2597–2649.

Eklund, J. and Kim, S. (2024). Fomc statement sentiment and inflation expectations. *Journal of Monetary Economics*, 141:45–62.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417.

Fleming, M. J. and Remolona, E. M. (1999). Price formation and liquidity in the us treasury market: The response to public information. *Journal of Finance*, 54(5):1901–1915.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–574.

Gorodnichenko, Y., Pham, T., and Talavera, O. (2023). The voice of monetary policy. *American Economic Review*, 113(2):548–584.

Grossman, S. J. and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408.

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.

Gürtler, M. and Gürtler, O. (2010). The effect of fomc statement language on financial markets. *Journal of Financial Research*, 33(4):369–391.

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.

Hayo, B. and Neuenkirch, M. (2010). Do federal reserve communications help predict federal funds target rate decisions? *Journal of Macroeconomics*, 32(4):1014–1024.

Hu, G. X., Pan, J., and Wang, J. (2013). Noise as information for illiquidity. *Journal of Finance*, 68(6):2341–2382.

Jarośinski, M. and Karádi, P. (2020). Deconstructing monetary policy surprises—the role of information shocks. *American Economic Journal: Macroeconomics*, 12(2):1–43.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1):161–182.

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. *Journal of Finance*, 66(1):35–65.

Lucca, D. O. and Moench, E. (2015). The pre-fomc announcement drift. *Journal of Finance*, 70(1):329–371.

Manela, A. and Moreira, A. (2017). News implied volatility and disaster concerns. *Journal of Financial Economics*, 123(1):137–162.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, 29(3):315–335.

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

Patton, A. J. and Verardo, M. (2012). Why do markets disagree? evidence from variation in opinion. *Review of Financial Studies*, 25(12):3734–3773.

Rosa, C. (2013). The high-frequency response of exchange rates and interest rates to macroeconomic announcements. *Journal of Banking & Finance*, 37(6):2162–2174.

Savor, P. and Wilson, M. (2014). Asset pricing: A tale of two days. *Journal of Financial Economics*, 113(2):171–201.

Schmeling, M. and Wagner, C. (2019). Does central bank tone move asset prices? *Review of Finance*, 23(5):933–972.

Shapiro, A. H., Sudhof, M., and Wilson, D. J. (2022). Taking the fed at its word: A new approach to estimating central bank objectives using text analysis. *Review of Economics and Statistics*, 104(4):768–784.

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.

Veldkamp, L. L. (2011). *Information Choice in Macroeconomics and Finance*. Princeton University Press.

Wang, K., Reimers, N., and Gurevych, I. (2021). Tsdae: Using transformer-based sequential denoising auto-encoder for unsupervised sentence embedding learning. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 671–688.

Wongswan, J. (2009). The response of global equity indexes to u.s. monetary policy announcements. *Journal of International Money and Finance*, 28(2):344–365.