

# **Analysis of the Relationship between Tone of the Fed Chair and Movement of the Bond Market**

By: Howard Bell

## **Introduction**

The monetary policy of the Federal Open Market Committee (FOMC) is a primary driver of the movement of the bond market and wider financial conditions. The Federal Reserve uses this policy to achieve their dual mandate of maximum employment and stable prices. The FOMC's actions influence investor's expectations for future economic conditions, borrowing costs, and more.

While the quantitative adjustments of the Fed through tools like the federal funds rate and QE/QT are central to traders, the communications of the Fed chair Jerome Powell at his post-meeting press conferences are equally so. His press conferences allow investors to glean insight into the Fed's current stance on monetary policy as well as their thinking about economic conditions and future policy paths. Investors interpret guidance on inflation, employment, and risk outlooks using the Fed chairs words. In fact, research shows that the market moves more than three times higher during press conferences given by Chair Jerome Powell than during press conferences by his predecessors Janet Yellen and Ben Bernanke. (Nairan)

With an FOMC meeting starting on January 27th, inflation running slightly higher than the Fed's target (around 2.7% as of December 2025) and the unemployment rate rising, hinting at a weakening labor market, fixed-income markets are evaluating both Fed actions and forward guidance for signs of future policy shifts.

With that in mind, I was left with two questions:

Can the tone and risk structure of Powell's press conferences be quantified and used to predict post-meeting Treasury price movements?

Can this be translated into a simple, interpretable trading rule?

## **FOMC Communication**

The FOMC holds eight scheduled meetings a year to discuss monetary policy decisions, around every six to eight weeks. Each meeting spans two days. On the first day, the twelve voting members review economic and financial data, hear forecasts, and discuss policy options. On the second day the members come together and decide on which monetary policy decisions to pursue. Their policy decisions are released in a statement at 2:00 p.m. ET, followed by a press conference with the Fed chair at 2:30 p.m which typically lasts around an hour. Jerome Powell's press conference is typically

split into two parts: a prepared opening statement which lasts around 10 minutes, then a much longer Q&A with reporters.

The press conference is the primary focus because it allows the Fed chair to add nuance and context to the committee's decisions. The written statement at 2:00 is technical and purposefully structured, while the press conference allows for follow-up questioning that can help markets delve deeper into the committee member's rationale. During the Q&A Powell may subtly reveal shifts in the economic outlook or policy opinions which were not fully captured in the statement. Subtle changes in wording and phrasing can be key to understanding the Fed's economic outlook. This makes the press conferences a good target for sentiment analysis as a machine-learning model may be able to detect patterns in language which cause certain movements in the market.

## **How my Approach Differs - Sentiment Analysis and FinBERT**

Fixed-income markets typically have expected rate cuts prior to the official announcement. For example, CME FedWatch data indicates a 95% probability that the Federal Reserve will keep interest rates unchanged in January 2026. As a result, prices only tend to move sharply if the policy statement or interest rate decision differs from consensus analyst expectations.

Rather than focusing on policy surprises alone, we will instead be using sentiment analysis to analyze and categorize Jerome Powell's words. Sentiment analysis is a NLP model used to identify and categorize the emotional tone behind a piece of text or interaction. In simpler terms it takes a sentence and produces a probability distribution across three sentiments: Positive, Neutral, and Negative. For each sentence it will give these three categories a rating from 0 to 1 continuous. The closer the value is to 1 the more confident the model is that the sentence aligns with the sentiment.

The model which we will be using, FinBERT, was trained on a variety of financial articles and market discussion, making it a viable tool for analyzing financial commentary like the Fed's monetary policy communications.

For greater clarity on how FinBERT may rank certain statements here are examples:

<b>Sentence</b>	<b>"We've had several labor market reports, which suggest, again, significant progress toward greater balance across a very—a broad range of indicators."</b>
Positive	0.9442650675773621
Neutral	0.04089755564928055
Negative	0.01483742706477642

FinBERT is strongly confident this sentence expresses positive financial sentiment, with about 94% confidence.

<b>Sentence</b>	<b>"I will have more to say about monetary policy after briefly reviewing economic developments."</b>
Positive	0.06718545407056808
Neutral	0.9057673811912537
Negative	0.027047188952565193

FinBERT is strongly confident that this sentence expresses a neutral financial sentiment with about 91% confidence.

<b>Sentence</b>	<b>"But now the hiring rate is very, very low"</b>
Positive	0.009403195232152939
Neutral	0.030794687569141388
Negative	0.9598020911216736

FinBERT is strongly confident that this sentence expresses a negative financial sentiment with about 96% confidence.

We will be using this model to score the sentiment of each of Jerome Powell's press conferences from 2023-2025 by splitting the transcript into smaller segments, scoring each segment with FinBERT, and combining those scores to produce meeting-level sentiment.

## Data Collection & Processing

We collected all labeled transcripts from the Federal Reserve website (<https://www.federalreserve.gov/>) spanning 2023-2025, keeping only dialogue from Jerome Powell and removing all speaker labels. Text was normalized by standardizing whitespace and splitting the transcript into individual sentences.

Each sentence was assigned a weight based on the presence of economically relevant keywords. Keywords focused on core policy ("inflation", "federal funds rate", "tightening" etc.) increased the weighting of the associated sentence to 2.0, while sentences containing other macroeconomic terms ("housing", "investment", "uncertainty", etc.) received a weight of 1.5. All other sentences were given a weight of 1.0. With these scores, a tone change metric was added to assess shifts in sentiment across consecutive meetings.

Using the sentiment probability distributions from FinBERT we then calculated weighted meeting tone as well as unweighted meeting tone. We also calculated negative and positive tail share, which is defined as the percentage of sentences which have a negative or positive probability distribution of greater than 60% respectively, capturing sentences where the model was very confident that the sentiment was skewed in one direction.

Finally, we added a caution ratio statistic, which calculates the ratio of confidently negative statements to confidently positive statements allowing one to more easily estimate skewed negative statements versus positive statements.

## Data Dictionary

Table 1 provides a comprehensive description of each field in the FOMC Meeting Sentiment Dataset.

Field Name	Data Type	Description	Example
meeting_mmddyy	Text (MM-DD-YY)	Date of FOMC press conference	12-10-25
n_sentences	Integer	Number of sentences analyzed in the meeting	365
weighted_tone	Float (Decimal)	Sentence-weighted average sentiment score for the meeting	-0.0351
unweighted_tone	Float (Decimal)	Simple average sentiment score across all sentences	-0.0422
neg_tail_share	Float (0–1)	Proportion of sentences with negative sentiment	0.1589

		probability > 60%	
pos_tail_share	Float (0–1)	Proportion of sentences with positive sentiment probability > 60%	0.0986
tone_std	Float (Decimal)	Standard deviation of sentence-level tone within the meeting	0.4353
caution_ratio	Float (Decimal)	Ratio of confidently negative to confidently positive sentences	1.61

Table 1: Data Dictionary for FOMC Meeting Sentiment Variables

## Basic Statistical Analysis

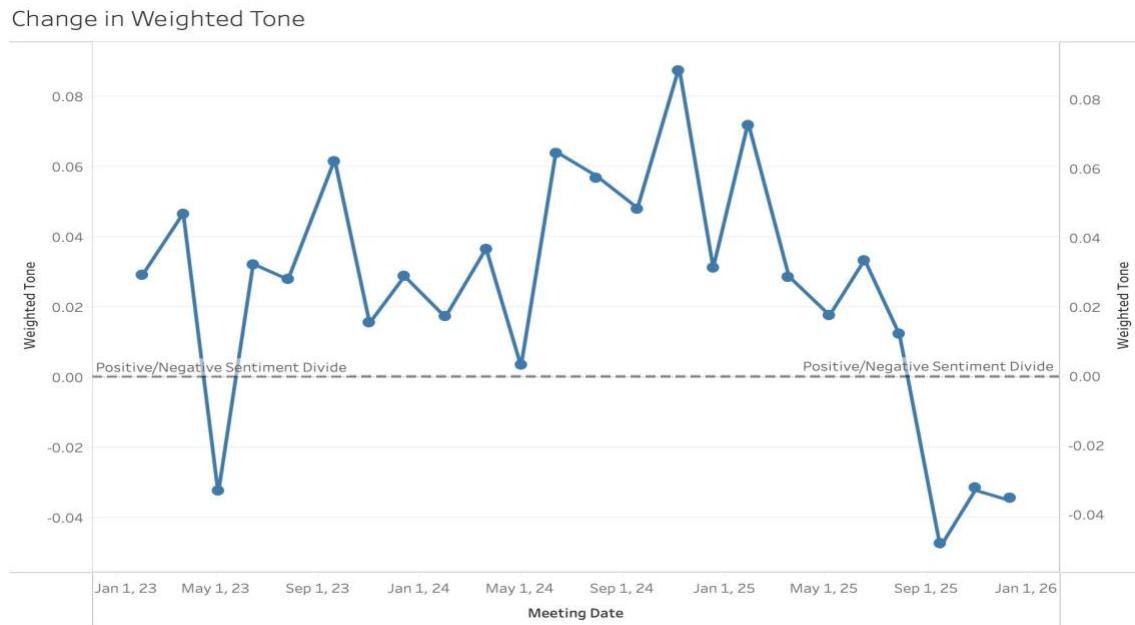
Jerome Powell's average sentiment across the sample period was slightly positive, with a mean weighted tone of 0.025. The model classified twenty of the twenty-four meetings as having positive sentiment. The lowest weighted tone occurred on September 17, 2025, coinciding with a weakening labor market, persistent inflation, and a 25-basis-point reduction in the federal funds target range. The highest weighted tone of 0.088 occurred on November 7, 2024 coinciding with inflation nearing the Fed's 2% target, cooling labor conditions as well as a 25-basis-point reduction.

Interestingly, FinBERT classified the last three meetings of 2025 in the sample as having a negative sentiment, a pattern not seen across any three meetings elsewhere in the dataset. This shift could be a combination of political pressure, a weakening labor market, and persistent inflation, which have led to a more pessimistic communication stance during press conferences.

The caution ratio had a positive correlation with weighted tone, indicating that meetings with more negative sentiment also tended to exhibit a higher concentration of negative language. However, there were exceptions like January 31, 2024 where the weighted tone was weighted positively (0.017) despite having a caution ratio exceeding one

(1.12). This meeting occurred during a period where economic performance was greater than expected but inflation still remained elevated.

Finally, the caution ratio is positively correlated with tone change in the following meeting ( $r = 0.38$ ,  $p = 0.075$ ). While this relationship does not meet the usual 5% significance threshold it does provide evidence that meetings with more negative sentiment are often followed by relative improvements in sentiment, showing mean reversion in Fed communication over time.



*Figure 1: Change in Weighted Tone Over Time*

## Trading Strategy

We can use daily price data from Yahoo Finance to examine how sentiment taken from Federal Reserve press conferences is associated with price movements in U.S. Treasury Bond ETFs. We will be analyzing the ETFs SGOV, GOVT, IEF, and TLT. Each of these ETFs track treasury securities with increasingly long maturities, going from 0-3 months (SGOV) up to 20+ years (TLT) allowing us to analyze effects across different durations of bond exposure.

To ensure that we properly capture the market response to Federal Reserve communication, we can align each FOMC meeting with the first trading day following the conclusion of the press conference to compute forward returns over a seven-day trading window. This ensures that we do not capture intra-day price movements during the event and instead only capture market reaction.

The linear regression was performed to connect post-conference ( $k = 7$ ) ETF returns with the sentiment measures calculated earlier, including weighted tone, positive tail

share, negative tail share, and the caution ratio. SGOV exhibits the strongest directional alignment with sentiment, weighted tone is positive ( $\beta = 0.001284$ ,  $p = 0.2388$ ,  $R^2 = 0.0625$ ), the caution ratio is negative ( $\beta = -0.000142$ ,  $p = 0.2119$ ,  $R^2 = 0.0699$ ), and the positive tail share remains positive in the tails-only model ( $\beta = 0.001893$ ,  $p = 0.1507$ ,  $R^2 = 0.1010$ ), but these effects are not statistically significant. For longer-duration ETFs (GOVT, IEF, TLT), explanatory power is near zero ( $R^2 \approx 0.0-0.08$ ) with insignificant coefficients, suggesting that any sentiment signal is concentrated at the short end.

These findings suggest that the press conference is most apparent at the short end of short-term Treasury instruments which are closely tied to upcoming policy expectations, though the coefficients are not statistically significant there is directional correlation. In contrast, longer duration ETFs (GOVT, IEF, TLT) show much weaker insignificant relationships demonstrating that longer-horizon yields are more majorly affected by broader macroeconomic factors.

With these results we can move to implementing and evaluating three sentiment-based trading strategies using caution ratio, weighted tone, and positive tail share to determine which is the most profitable strategy in the sample. For weighted tone and positive tail, we will be implementing a simple strategy where if the variable is above its median value we will go long on SGOV at the beginning of the next trading day and sell after 7 days. For the caution ratio we will go short when it is above its median and sell after 7 days. To ensure consistency with the regression framework, each strategy enters positions at the close of the first trading day following an FOMC press conference and holds the position for seven trading days. We will evaluate these strategies using a starting portfolio value of \$100,000.

## **Baseline**

To benchmark our strategy, we should first evaluate with a baseline. For the baseline scenario we will go long SGOV following every FOMC meeting, reflecting the return an investor would earn by passively holding short-duration Treasuries during post-meeting periods. this baseline executes 24 trades and produces an average return of 0.127% per trade (0.001270), a 100% hit rate, and a cumulative return of 3.092% (0.030923).

## **Weighted Tone Strategy**

The strategy which went long on SGOV when the weighted tone was greater than the median participated in 12 trades, had an average return per trade of 0.1315%, a hit rate of 100%, and a cumulative return of 1.589%.

## **Positive Tail Strategy**

The strategy which went long on SGOV when the positive tail was greater than the median participated in 12 trades, had an average return per trade of 0.1334%, a hit rate of 100%, and a cumulative return of 1.612%.

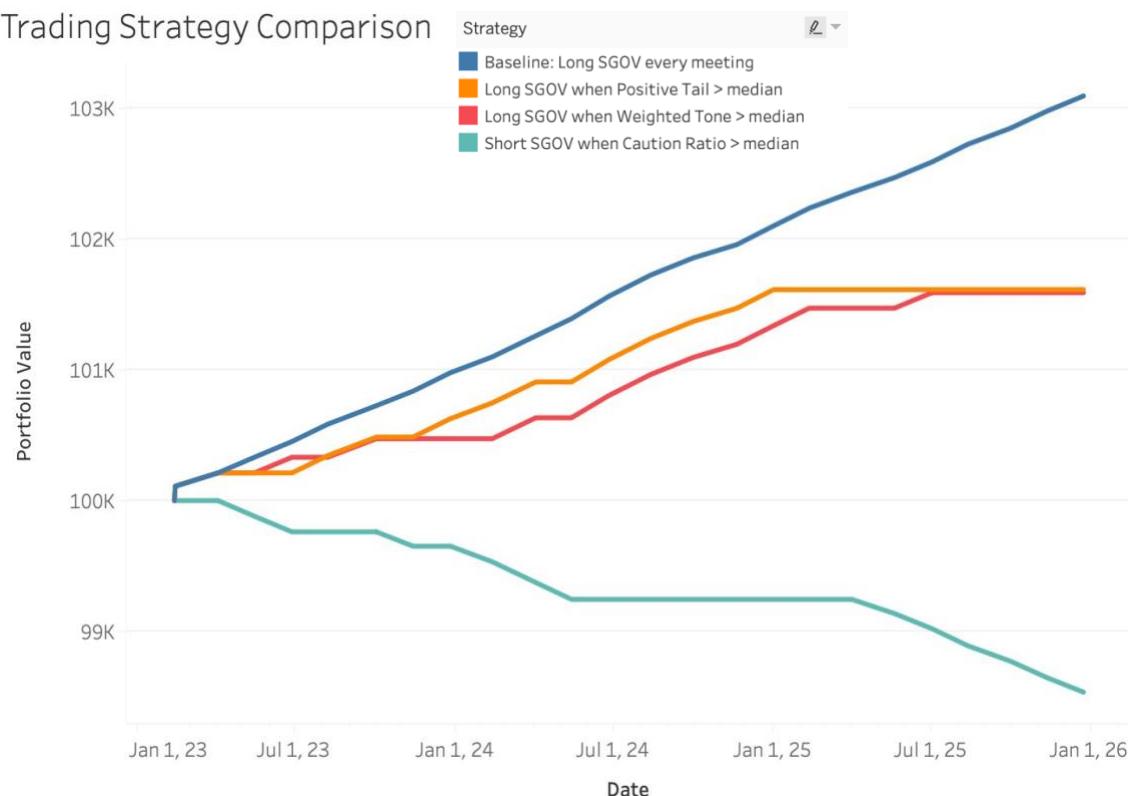
## **Caution Ratio Strategy**

The strategy which went short on SGOV when the caution ratio was above its median participated in 12 trades and had an average return per trade of  $-0.1228\%$  with a cumulative return of  $-1.464\%$ .

The two long strategies generate positive average returns per trade and perfect hit rates within the sample, reflecting very selective exposure. However, these strategies trade less frequently than the baseline, resulting in lower cumulative returns relative to the that passive trading strategy. The short strategy based on the caution ratio performs poorly, generating negative returns and a low hit rate, suggesting that strongly cautious language does not reliably predict downward moves in short-term Treasury prices.

Although sentiment-based strategies appear to limit downside risk, they do not outperform the baseline strategy. The majority of positive returns could be attributed to SGOV's structural carry rather than to sentiment-driven price movements.

These results indicate that Federal Reserve sentiment contains economically meaningful information about short-term interest rate expectations, but that this information is not sufficient to generate excess returns over a passive allocation of short-term bond markets.



*Figure 2: Trading Strategy Comparison*

## Conclusion

This study shows that sentiment extracted from Federal Reserve press conferences is most strongly associated with short-duration Treasury returns. Positive tone and confidently positive language are linked to higher SGOV returns, while caution is associated with weaker short-term performance, with these effects reducing at longer maturities. Although sentiment-based strategies generate positive in-sample returns, they do not outperform a passive post-meeting allocation, suggesting that Federal Reserve communication primarily shapes market expectations rather than producing persistent trading opportunities.

## Limitations and Future Work

This analysis has several limitations that could potentially be expanded upon in the future:

- Sample size of 24 meetings is relatively small; extending to 2018-2025 analysis window would provide more statistical power
- Strategy performance is tested on the same data used to calculate median thresholds, leading to potential overfitting
- This analysis ignores transaction costs which would reduce returns in a real-world trading scenario
- The analysis was almost entirely focused on short-term treasuries where analysis of different assets could produce more significant price movements

Future research could explore:

- Combining sentiment with other variables (market volatility, surprise, etc.)
- Real-time sentiment analysis using live AI transcription
- Larger training set which permits a training/testing dataset split

## Citations

Narain, Namrata, and Kunal Sangani. "The Market Impact of Fed Communications: The Role of the Press Conference." International Journal of Central Banking, forthcoming, <https://www.ijcb.org/forthcoming-articles/market-impact-fed-communications-role-press-conference>