A Modern Exploration of the Impact of Monetary Policy Statements on Asset Price Dynamics

Eytan Rozenblum
er526@cornell.edu

Berend van Nieuwland
bnv7@cornell.edu

Abstract

This study revisits the seminal findings of Gurkaynak et al. [2005] on the factor structure underlying asset price responses to Federal Open Market Committee (FOMC) announcements, extending their analysis through 2024. We replicate their methodology using updated data and confirm the presence of at least two distinct dimensions describing how financial markets react to monetary policy announcements. While one factor is broadly consistent with the immediate surprise in interest rate decisions, we find that the correlation between this factor and conventional measures of policy surprise is much weaker in the modern era than in their original analysis.

To gain deeper insight, we introduce a new "tone surprise" measure, derived from a recent hawkish-dovish classification dataset of FOMC statements. This measure captures unexpected shifts in the Federal Reserve's communicative stance. Our results suggest that the tone surprise measure better aligns with observed asset price reactions under certain market regimes — specifically, during periods of heightened interest rate volatility — than do traditional measures of immediate target rate surprise.

Overall, our evidence highlights that while the fundamental two-factor framework remains robust, the nature of these factors has evolved due to changes in the Federal Reserve's communication strategies and market structures. Accounting for the nuanced and dynamic interpretation of forward guidance and tone offers a richer understanding of how monetary policy announcements shape asset prices in the modern financial landscape.

1 Introduction

Monetary policy plays a central role in shaping economic activity, influencing everything from inflation and unemployment to asset prices and market stability. Researchers have long tried to model the mechanisms through which monetary policy spreads into the economy, namely focusing their efforts on understanding how monetary announcements impact financial markets. A landmark contribution to this field was *"Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements"* (Gurkaynak et al. [2005]). This seminal work challenged the prevailing belief that the effects of US monetary policy on asset prices could be captured by a single factor: changes in the federal funds target rate. Instead, the authors identified a second factor that captures the expected future path of monetary policy, fundamentally reshaping how economists and market participants interpret monetary policy announcements.

The core insight of Gürkaynak et al. was derived from analyzing the response of asset prices (S&P 500, T-bills, T-notes) in a tight time window (thirty minutes to one hour) surrounding Federal Open Market Committee (FOMC) statement releases. Their findings highlighted the dual impact of monetary policy: the immediate effect of rate changes and the longer-term expectations about future policy.

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

This research, although significant, may be outdated. Monetary policy announcements by the FED have shifted in philosophy between the paper's data (1990-2004) and ours (2004-2024). Central bankers used to make announcements in a more obscure way than their modern counterparts (as Greenspan famously said to the Congress: *"Since I've become a central banker, I've learned to mumble with great incoherence. If I seem unduly clear to you, you must have misunderstood what I said,"*. They have since implemented structural changes with forward guidance, to tell the public about the likely future course of monetary policy. The financial data available to the public (exemplified by the growth of Bloomberg LP), and the democratization of high-performing computing machines, have also revolutionized financial analysis (as is exemplified by the 'quant' revolution in modern market finance). For these reasons, we wanted to understand whether the insights from Gürkaynak et al. were still valid today and, if not, how we could update them. Our research aims at answering the question: **how can we explain the response of asset prices to monetary policy in the US?**

Consequently, the goal of this study is threefold:

- **Replication and Validation**: To replicate the original findings of Gurkaynak et al. using updated data, assessing whether the two-factor framework continues to explain market responses to FOMC announcements.
- **Exploration of Alternative Methods**: To apply Independent Component Analysis (ICA), a statistical technique designed to identify independent sources of variation, as an alternative approach to interpreting monetary policy factors.
- **Integration of Modern Methods:** Given the importance that the way the speeches are delivered has taken in the past decades (thanks to forward guidance), to enrich the analysis by integrating sentiment and tone analysis data from FOMC statements. This will allow us to examine whether the tone of monetary policy communication influences asset prices and, if so, if it is a better explanation of the main factor than the ones found by Gürkaynak et al.

By addressing these objectives, this project not only seeks to validate the robustness of the original findings, but also to provide a more in-depth and comprehensive understanding of how monetary policy influences financial markets in the modern era. Ultimately, this study contributes to the ongoing discourse on the intersection of central bank communication, market expectations, and asset pricing.

2 Datasets Overview

In this section we will discuss the datasets used and the respective cleaning and preprocessing decisions, while ensuring results are predictable, computable, and stable.

2.1 Financial Data and FOMC Meeting Dates

To assess the impact of the announcement on asset prices, Gürkaynak et al. (2005) use high frequency data (30-minutes and 1-hour window around the announcement) of the relevant assets for which they want to understand the impact of the announcement, for all FOMC announcements from January 1990 through December 2004 (133 observations). They use data of treasury bills and notes (three-month, six-month, two-year, five-year, and ten-year yields), futures contracts on bonds (current-month and three-month-ahead federal funds futures contracts), futures contracts on forex (two-, three-, and four-quarter-ahead eurodollar futures contracts), and equities (S&P500).

Comparing the results that they obtain with intraday data and with daily data, they find that *"the surprise component of monetary policy announcements can be measured very well using just daily data"* and that *"the [. . .] intraday measures are quite similar to the daily measure"*. For this reason, we used daily data for all the assets which we track: treasury bonds (three-month, six-month, twelve-month, three-year, five-year, and ten-year yields), spot forex prices (eurodollar), and equities (S&P500 and NASDAQ). We have data for all FOMC announcements from January 2004 to March 2024 (187 observations). Our data is extracted from Bloomberg and LSEG Data & Analytics (formerly Refinitiv).

We found that these data sources were very reliable for our purpose, as there was little-to-none data missing, and the data provided was in line with cross-checks from other online datasources.

Our data is in form of multiple time series, with one observation at the end of each day. To get the variation of each asset around the time of the FOMC announcement, we scrape the website of the

Federal Reserve for the dates of the meetings from 2004 to 2024 using `BeautifulSoup` and create a database with the variation of our asset of interest around the time of the meeting (variation between the last closing price before the meeting and the first one after the meeting).

2.2 A Trillion Dollar Words: A New Financial Dataset

This project utilizes a specialized dataset derived from *"A Trillion Dollar Words: A New Financial Dataset"* (Shah et al. [2023]). The dataset, which spans 1996–2022, was created to provide a hawkish-dovish classification of FOMC statements. It is built upon a rigorous methodology to ensure its utility for analyzing monetary policy's nuanced impacts on financial markets.

The dataset comprises manually labeled sentences from FOMC statements, including meeting minutes, speeches, and press conference transcripts. Each sentence is classified on a "Hawkish-Dovish-Neutral" scale based on its monetary policy stance.

- **Dovish Sentences**: Indicate future monetary policy easing, such as lowering interest rates or stimulating economic growth.
- **Hawkish Sentences**: Indicate future monetary policy tightening, such as raising interest rates or controlling inflation.
- **Neutral Sentences**: Either have mixed sentiments indicating no significant change or are unrelated to monetary policy stance.

To ensure high-quality and unbiased labeling, the curators of the dataset have ensured a rigorous sampling design. Firstly, the annotation process involved two annotators independently labeling sentences using a predefined annotation guide. After the initial labeling, the results were compared, and any inconsistencies were resolved to maintain alignment and accuracy. To address the Fed's practice of using mixed messaging to mitigate excessive market reactions, a custom sentence-splitting method was developed. This method identified and split sentences containing contrasting stances at specific keywords, such as "but," "however," and "although," enabling the isolation and accurate classification of distinct tones within a single statement [Shah et al., 2023].

Finally, the classified sentences were aggregated into 214 document-level measures to quantify the overall monetary policy tone on specific FOMC release dates. This measure, reflects the relative frequency of hawkish versus dovish sentences, normalized by the total sentence count in the document. For a statement released at time t, its hawkish-dovish score is given by:

$$HD_t = \frac{\#\text{Hawkish}_t - \#\text{Dovish}_t}{\#\text{Total}_t} \tag{1}$$

2.3 Federal Reserve Bank of St-Louis

The additional dataset this project utilizes is economic data from the Federal Reserve Bank of St. Louis's FRED database, spanning from 1996 to 2024. The dataset includes key macroeconomic indicators relevant to monetary policy analysis, as detailed in Table 1.

Table 1: Overview of Features

Feature	Description
GDP	Gross Domestic Product
MORTGAGE30US	30-Year Fixed Mortgage Rate in the US
PCEPI	Personal Consumption Expenditures Price Index
UNRATE	Unemployment Rate

The data on the variables retrieved from FRED's database is clean and free of missing values, so no imputation methods were necessary. All variables are on a monthly-level basis, except for GDP as this was only available on a quarterly-level. The goal is to merge this dataset with the Trillion Dollar Words dataset, and align it with the FOMC statement release dates. For that reason, the monthly-level data has been aggregated and averaged in between FOMC release statements. For the GDP variable, the closest observation before the release of the FOMC statement was used.

Finally, this dataset is merged with the Trillion Dollar Words dataset and indexed according to the release date of the FOMC statements, resulting in a dataset with 214 entries ready for analysis in a predictable, computable and stable framework.

3 Explanatory Data Analysis

3.1 Multicollinearity

We noticed an extremely high correlation between PCEPI and GDP (as seen in Figure 1 in Appendix), in the correlation matrix of the features. This could indicate multicollinearity, which can introduce stability and predictability issues in the modeling of our problem. To further investigate this, the Variance Inflation Factor (VIF) between the variables was calculated, with results reported in Table 3 in Appendix.

The table reports extremely high VIF values, indicating that multicollinearity is indeed present. This needs to be accounted, and will be discussed in the next section.

3.2 Feature Engineering

Based on the multicollinearity analysis, we observed strong correlations among variables such as `GDP` and `PCEPI`. To mitigate this, we performed feature engineering and reducing dimensionality where necessary. Subsequently, we constructed two new variables: `GDP_GROWTH` and $\mathtt{HD}_{t-1}$. `GDP_GROWTH` was engineered by computing the growth between two subsequent periods, and $\mathtt{HD}_{t-1}$ indicates the lagged variable of $\mathtt{HD}_t$. As a verification, we computed the VIF for our finally settled on model, resulting in the use of `GDP_GROWTH`, `UNRATE` and $\mathtt{HD}_{t-1}$, with healthy VIF values, which may be seen in Table 4 in the Appendix.

4 Methodology

4.1 Number of Dimensions of Markets' Response to Monetary Policy Announcements

Let X denote the $T \times n$ matrix, with rows corresponding to monetary policy announcements, columns corresponding to asset prices, and each element of X reporting the change in the corresponding asset price before and after the corresponding announcement. Following Gurkaynak et al. [2005], we define:

$$X = F\Lambda + \eta \tag{2}$$

Where F is a $T \times k$ matrix of unobserved factors (with $k < n$), Λ is a $k \times n$ matrix of factor loadings, and η is a $T \times n$ matrix of white noise disturbances. We wish to determine k, i.e., the number of factors required to adequately describe X. For this, we use the matrix rank test of Cragg and Donald [1997], which tests the null hypothesis $\mathcal{H}_0$: "X is described by k_0 common factors" against $\mathcal{H}_1$: "X is described by $k > k_0$ common factors."

We define:

$$\min_{\Lambda \in \mathcal{M}_{k_0 \times n}(R)} \left((\Sigma_{k_0})^T \, \hat{\Omega} \Sigma_{k_0} \right) \tag{3}$$

Where:

$$\Sigma_{k_0} = \text{vech}(\Sigma_X) - \text{vech}(\Sigma_{F\Lambda_{k_0}+\eta})$$
$$\hat{\Omega} = \text{Cov}(\text{vech}(\Sigma_X))$$

Under $\mathcal{H}_0$ This is a Wald test with a limiting distribution of χ^2 with $\frac{n(n+1)}{2} - (n \cdot k_0) + \frac{k_0(k_0-1)}{2}$ degrees of freedom.

Following Gurkaynak et al. [2005], we then estimate the unobserved factor matrix F using the standard method of principal components applied to our data matrix X. This is a dimensionality reduction algorithm that projects data along principal components, uncorrelated variables constructed as linear combinations of the initial variables.

4.2 Independent Component Analysis

To get an initial interpretation of the factors that we uncover, we use the Independent Component Analysis framework (implemented by[Comon, 1994]). It provides an initial decomposition of the factors influencing asset price movements following FOMC statements. While traditional factor analysis identifies correlated components, ICA seeks statistically independent components, potentially revealing new insights into the drivers of asset price responses. This analysis allows us to understand the component that influence each of the factors describing the response of asset prices to monetary policy, giving an initial interpretation for these.

A key assumption of ICA is that the features are non-gaussian as it uses that characteristic to separate the signals. This is validated as we are working with financial data for this method. ICA is generally well-suited for analyzing financial time-series data where hidden independent sources may influence observable variables.

The ICA model is defined as follows:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{4}$$

Where $\mathbf{x}$ are the observed features, $\mathbf{A}$ is a mixing matrix, and $\mathbf{s}$ are the latent independent components. Rearranging for estimation, this results in:

$$\hat{\mathbf{s}} = \widehat{\mathbf{W}}\mathbf{x} \tag{5}$$

Where $\widehat{\mathbf{W}} = \widehat{\mathbf{A}}^{-1}$, the estimated unmixing matrix, $\hat{\mathbf{s}}$ are the estimated independent components, and $\mathbf{x}$ the observed features. The unmixing matrix, $\mathbf{W}$, is estimated by maximizing statistical independence and non-gaussianity, using the FastICA algorithm [Hyvärinen and Oja, 2000]

4.3 Structural Interpretation of the Factors

To give a structural interpretation to the k_0 factors that they find, Gurkaynak et al. [2005] rotate the matrix F into a matrix Z ($Z = FU$) with two columns which are still orthogonal, but for which the second factor has no effect on the current federal funds rate. In other words, the rotation is made so that the second column of Z is a vector that is associated on average with no change in the current-month federal futures rate.

As a result, they interpret the unexpected change in the current target for the federal funds rate as being driven exclusively by Z_1 (plus some white noise). In this setting, Z_1 explains the surprise of the market to the announcement for the current rate, whereas Z_2 includes any information (besides the decision for the current target rate) that affects the expected path for monetary policy over the upcoming year. For this reason, they refer to Z_1 and Z_2 as the "target" factor and the "path" factor.

The target factor Z_1 defined in this way should be similar to a measure of federal funds target surprises which the authors introduce. They define:

$$Policy\ Surprise_t = \mathbb{E}_t(r_t) - \mathbb{E}_{t-\Delta t}(r_t) \tag{6}$$

Where r_t is the rate announced at time t by the FED.

To quantify this measure of policy surprise, the authors build on the insights of Kuttner (2001): for a given time τ, the price of the FED's 30 days one-month futures, ff_τ, is the average of the rate that has prevailed in the month so far and the rate that is expected to prevail for the end of the month, weighted for the number of days for which each rate prevails (plus a measure ρ of the risk premium of the contract).

$$ff_\tau = \frac{d}{D} \cdot r_{t-1} + \frac{D-d}{D} \cdot E_\tau(r_t) + \rho_\tau \tag{7}$$

where d is the day of the month, D is the total number of days in the month, r_{t-1} is the rate that has prevailed so far in the month, $E_\tau(r_t)$ is the market expectation at time τ of the rate that will prevail for the rest of the month, and ρ_τ is the risk premium of the contract. Therefore, for Δt small enough so that $\rho_t - \rho_{t-\Delta t}$ is negligible (i.e., the market does not have the time to price a new premium for the futures contract), we immediately derive a measure for Equation (6):

$$Policy\ Surprise_t = (ff_t - ff_{t-\Delta t}) \cdot \frac{D}{D-d} \tag{8}$$

4.4 A New Measure for Market Surprise

To provide a deeper analysis and enhanced modeling stability, we propose a new method for estimating the market surprise, using the tone of the announcements. We define `tone surprise` as the difference between the observed tone of the FOMC released statement and what the market expects the tone of that same statement to be:

$$Tone\ Surprise_t = \mathbb{E}_t[HD_t] - \mathbb{E}_{t-\Delta t}[HD_t] \tag{9}$$

The first term in Equation (9) is observed in the Trillion Dollar Words dataset. The second term, representing the market's expectation of the tone, is unobserved and must be estimated. To do this, we split the HD index provided by the Trillion Dollar Words dataset into a training and test dataset and regress the training dataset on a set of macroeconomic indicators available at $t - \Delta t$ — indicators accessible to the market and which it assumes that the FED uses to guide its decisions. These serve as a proxy to quantify the anticipated market expectation of the FED's tone. Consequently, the model is trained using macroeconomic indicators as regressors and previously observed hawkish-dovish measures as the target variable, in order to estimate the market's expectation of the tone. Through this method we can estimate `Tone Surprise` by looking at the difference of the observed value and the proxy, resulting in Equation (10) reported below.

$$\widehat{Tone\ Surprise}_t = HD_t - \widehat{\mathbb{E}}_{t-\Delta t}[HD_t] \tag{10}$$

Where HD_t is the observed hawkish-dovish tone at time t and $\widehat{\mathbb{E}}_{t-\Delta t}[HD_t]$ is the estimation of what the market expects, just before the release, of what the tone of that FOMC announcement will be.

5 Results

5.1 Number of Dimensions of Markets' Response to Monetary Policy Announcements

Using the methodology presented in subsection 4.1, we applied the matrix rank test to our full rank dataset of financial assets, comprising 9 columns: treasury bonds (three-month, six-month, twelve-month, three-year, five-year, and ten-year yields), spot forex prices (eurodollar), and equities (S&P500 and NASDAQ).

The results of this regression are presented in Table 2. We find that the hypothesis that the response of asset prices to monetary policy announcements is characterized by zero or by one common factor is clearly rejected (p-value $<< 1\%$). This validates the initial intuition by Gürkaynak, Sack, and Swanson: a single factor (namely, the surprise factor in changes in the federal funds rate) is not sufficient to adequately characterize changes in asset prices. These results also show that we cannot reject $\mathcal{H}_0$: "X is described by 2 common factors" at a 95% level of confidence.

Table 2: Summary of Rank Test Results for our Complete Dataset

Dataset	Rank	Min Distance	Deg of Freedom	Critical Value	p-Value
Full Dataset	0	121.89	36	51.00	0.00
Bonds, Forex,	1	59.60	27	40.11	0.00
Equities	**2**	**29.83**	**19**	**30.14**	**0.05**
(9 columns)	3	16.20	12	21.03	0.18

As a stability test, we conducted the rank test on three sub-datasets of this original dataset, with results reported in Table 5 in Appendix. The first dataset includes all treasury bonds as well as spot Eurodollar, the second also includes stock prices for the S&P500, the third includes stock prices for the NASDAQ instead of the S&P500. These financial assets have different answers to a change in target rate. For instance, the United States' largest stocks (S&P500) may have a different reaction than tech stocks (NASDAQ), which are capital-intensive and intensively use debt, therefore being more sensitive to FED rates. This is why this is a good stability test.

These results back the ones from Table 2. We also find that they validate the hypothesis $\mathcal{H}_0$ more strongly than the results provided by Gurkaynak et al. [2005], because we use matrices of higher maximal rank (respectively 7, 8, 8, and 9) than they do (the matrices on which they run the test have 6, 5, and 5 columns, respectively).

For subsequent parts, we focus on our most comprehensive dataset, which includes treasury bonds (three-month, six-month, twelve-month, three-year, five-year, and ten-year yields), eurodollars, and aggregate equities (S&P500 and NASDAQ) as it is the most representative of the economy's reaction to Fed announcements.

5.2 Independent Component Analysis

The results from the ICA implementation are displayed in Figure 2 in Appendix. From the Figure we see that the first factor loads heavily on indices/equities and tends to load more heavily towards long-term bonds, which are treasury notes, denoted `'gt'`. The second factor is more loaded towards short-term bonds, namely treasury bills, denoted `'gb'`, and barely loads on indices. As a stability test, an implementation with the 3-year bond feature dropped, was performed, resulting in the same results (see Figure 3 in Appendix).

These results suggest that one factor (IC1) captures a more general "risk sentiment" or "market-wide" response to FOMC releases, as it heavily loads on equities and longer-term maturities, reflecting how longer horizon interest rate expectations and growth prospects influence stock prices and longer-term borrowing costs. In contrast, the second factor (IC2), which is most pronounced in short-term bond yields and does not significantly affect stock indices, likely represents the market's immediate reaction to the Fed's near-term policy stance and changes in the expected path of short-term interest rates.

This initial analysis corroborates the results from Gurkaynak et al. [2005]. One factor impacts more the bonds with maturity < 1 year (what Gurkaynak et al. [2005] analyze as the announcement factor) and the other impacts the bonds with maturity > 1 year more. That is, in the short-to-medium term, words do speak louder than actions, whereas in the long end of the yield curve actions are more relevant than words, and the announcement factor is not significant to predict changes in asset prices. Nevertheless, despite these overarching behaviors, we see that both factors load on medium-to-long term treasury notes, hinting at the fact that the structural interpretation of the factors may not be as clear as the one from Gurkaynak et al. [2005].

5.3 Structural Interpretation of the Two Factors

Gurkaynak et al. [2005] find that Z_1 is very close to the vector measuring market surprise , with "a correlation of over 95 percent". However, in our analysis, we find very low correlation (approximately 4%), suggesting that Z_1 cannot be interpreted as a target factor that captures well the markets' surprise to policy announcements. This discrepancy hints at potential differences in market dynamics or data interpretations in our study compared to the original findings.

We impute this result on two main factors. Primarily, the FED changed the way it announces its decisions between the paper's period of interest (1990 to 2004) and ours (2004 to 2024), relying more heavily on forward guidance in the announcements since the turn of the century. Also, our period of interest was exceptional in terms of monetary policy (they were either abnormally low or extremely volatile, see Figure 4 in Appendix).

This insight drove our next analysis. First, we wanted to analyze whether the market's answer to the FED's announcements could be better modeled using a measure of surprise to the tone of the announcement, rather than the announcement itself. Second, we wanted to understand whether we could give a better interpretation of both factors by controlling on the macro monetary environment (i.e., by controlling on the level of the federal funds rate).

5.4 A New Measure for Market Surprise

5.4.1 Computation of the New Measure for Market Surprise

The computation of $\widehat{Tone\ Surprise}_t$ in Equation (10) requires modeling choices.

- **Train/ test split**: We have hawkish-dovish measures from 1996 to 2022. What kind of split between train and test data makes the most sense?
  - Random split to ensure that our training data is not overly influenced by the tone of one FED chair while our test data only includes announcements from other chairs.
  - Temporal split (before the date for which we have Bloomberg data — 2004 — vs. after) so that we can compare our tone surprise measure vs. Gurkaynak et al. [2005] policy surprise measure on the same announcements.
  - A mixture of both to mitigate the drawbacks from both methods.
- **Regressors**: We have multiple regressors from the Federal Reserve Bank of St-Louis, as detailed in Table 1. Which mixture of regressors make the most sense? (i.e., Which regressors do market actors think are the most important for the FED when drafting their policy statement?)
- **Regression Method**: What makes most sense between OLS linear regression, ridge, lasso, or random forest?

We opted to conduct a ridge regression on unemployment rate, GDP growth, and HD_{t-1}, with a temporal split between train and test. This is because we wanted a regularization term in our regression while avoiding the sparsity in regression coefficients that LASSO can induce. Using these regressors came from our domain knowledge. Finally, using this train/ test split made the most sense to ensure that we could evaluate the performance of our tone surprise measure against the policy surprise from Gurkaynak et al. [2005] on the same set of announcements.

This regression yielded an RMSE of 0.1348 and an R^2 of 67.50%, indicating good generalization performance on the test data.

5.4.2 PCS Testing for this New Measure for Market Surprise

As a stability test, we performed this regression with all possible configurations discussed at the beginning of this subsection. We found that the best models had comparable performance to the one we selected (when measuring them on RMSE and R^2).

Also, we expect such a financial signal to be normally distributed. A good prediction of the market's surprise to the tone of the FED's announcement should therefore have normal distribution. Therefore, as a predictability test, we tested normality of the signal which we obtained. Performing a permutation test with 100,000 permutations for the Kolmogorov-Smirnoff statistic to test $\mathcal{H}_0$: "The distribution is normally distributed" against $\mathcal{H}_1$: "The distribution is not normally distributed", we find that we cannot reject $\mathcal{H}_0$ at any usual level of confidence (p-value » 10% for all the permutation tests that we ran). The frequency distribution of the tone surprise with a fitted Gaussian curve is represented in Figure 5 in Appendix, and the results of the permutation test with the Kolmogorov-Smirnoff statistic are in Figure 6 in Appendix.

5.5 Update on the Interpretation of Gürkaynak et al.

After conducting the same rotation that was described in subsection 4.3 and applied in subsection 5.3, we find a correlation of 10% between the first factor of the rotated matrix and the tone surprise measure. This is better than the previous results on the correlation between the policy surprise and the first factor of the adequately rotated matrix, but it is far from the results presented in Gurkaynak et al. [2005].

Because our period of interest is particularly exceptional in terms of monetary policy, with rates that were either abnormally low or extremely volatile (see Figure 4 in Appendix), we controlled for the regime of monetary policy ("low rates" and "volatile rates"). Results are reported in Figure 7 in Appendix.

- In periods of low rates, the first component (after adequate rotation) is most correlated with the measure of policy surprise (21.0%, 5 times more than without controlling for rates). Even after adequate rotation, it is only correlated up to 3.5% to the tone surprise measure. This hints at the fact that market agents may have expected rates to take off faster than they actually did during the zero-rate period (2009-2015 and 2020-2022), with higher market corrections when the rates remained flat.
- In periods of high volatility, the first component (after adequate rotation) is most correlated with the measure of tone surprise (44.1%, 4.5 times more than without controlling for rates). It is only correlated up to 4.3% to the measure of policy surprise. This indicates that agents may have high expectations from the forward guidance during times of high volatility.

During these periods, the market surprise measure that does not closely align with asset price responses to monetary policy isn't stagnant or zero; it simply evolves in a different pattern than asset prices do. The interpretation provided above reflects the varying emphasis placed on short-term policy ("actions") versus medium- to long-term announcements ("words") throughout different phases of the monetary policy cycle.

6 Conclusion and Discussion

Our findings confirm that the reaction of asset prices to FOMC announcements continues to be characterized by at least two latent dimensions, affirming Gurkaynak et al. [2005]'s fundamental insight that a single-factor explanation is insufficient. However, the interpretation of these factors in the post-2004 era appears to have shifted. Traditional measures of immediate policy surprise, closely aligned with the short-term interest rate decision, no longer maintain the strong correlation with market responses that they once did. Instead, our results indicate that forward guidance and the tone of monetary policy communications — elements that have become more transparent and data-driven — play a more pronounced role in shaping asset price movements.

By introducing a tone surprise measure derived from a hawkish-dovish classification of FOMC statements, we provide a new lens through which to interpret the evolving relationship between central bank communications and financial markets. While this measure does not universally dominate the explanatory power of the traditional target-rate-based surprise, it becomes particularly salient during periods of elevated uncertainty and volatility. In such regimes, the market appears to place greater weight on the central bank's forward-looking narrative and sentiment, rather than on the immediate rate setting alone.

These insights have several implications. First, it validates that central bankers' communications — both in terms of forward guidance and narrative tone — are integral policy tools that can significantly influence market outcomes. Second, researchers and practitioners may benefit from integrating text-based sentiment measures and machine learning techniques into their analytical frameworks. Doing so could offer a more comprehensive understanding of monetary policy's transmission mechanisms, especially under nontraditional monetary environments and complex market conditions.

Finally, our study underscores that the dynamics of monetary policy communication and market response are continuously evolving. Policymakers and market participants alike must remain attentive to these shifts. Updating factors, metrics, and models—just as we have done here—ensures ongoing relevance in an era of rapid changes in central banking communication, financial data availability, and computational methods.

7 Limitations and Next Steps

While this study provides valuable insights into the relationship between monetary policy communication and asset prices, it has limitations that present opportunities for future work. Firstly, the analysis does not fully account for temporal dynamics. Employing models like ARIMA or ARIMAX in future research could better capture time-dependent patterns and assess their impact on the results. Secondly, the dataset is inherently small, constrained by the number of FOMC statements, and future studies could benefit from a larger dataset as more data becomes available. Lastly, the focus of this study is exclusively on the US market, aligning with the original work of Gurkaynak et al. [2005]. Future research could explore the applicability of these findings to other regions (ECB, BoJ, etc.) to assess whether similar patterns emerge in different economic contexts.

References

Charles W. Calomiris and Harry Mamaysky. How natural language processing will improve central banks' accountability and policy. *Cato Journal*, 40(2):327–347, 2020. URL `https://www.cato.org/cato-journal/spring/summer-2020/how-natural-language-processing-will-improve-central-bank`.

Pierre Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3): 287–314, 1994. Retrieved from `https://www.sciencedirect.com/science/article/pii/0165168494900299`.

John G. Cragg and Stephen G. Donald. Inferring the rank of a matrix. *Journal of Econometrics*, 76 (1):223–250, 1997. ISSN 0304-4076. doi: https://doi.org/10.1016/0304-4076(95)01790-9. URL `https://www.sciencedirect.com/science/article/pii/0304407695017909`.

Refet S. Gurkaynak, Brian P. Sack, and Eric T. Swanson. 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, 2005. Retrieved from `https://www.ijcb.org/journal/ijcb05q2a2.htm`.

Gábor Gáti and Jake Handlan. Monetary communication rules. ECB Working Paper Series 2759, European Central Bank, 2023. URL `https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp2759~085f242f95.en.pdf`.

Stephen Hansen, Michael McMahon, and Andrea Prat. Transparency and deliberation within the fomc: A computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870, 2018.

Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural Networks*, 13(4–5):411–430, 2000. doi: 10.1016/S0893-6080(00)00026-5. Received 28 March 2000; accepted 28 March 2000.

Aniket Shah, Sandeep Paturi, and Sudheer Chava. Trillion dollar words: A new financial dataset, task & market analysis. *arXiv preprint*, 2023. arXiv:2305.07972. Retrieved from `https://arxiv.org/abs/2305.07972`.

Bin Yu and Rebecca L. Barter. *Veridical Data Science: The Practice of Responsible Data Analysis and Decision Making*. The MIT Press, 2024. Retrieved from `https://vdsbook.com/#use-of-code`.

A Appendix and Supplemental Material

A.1 Literature Review

A.1.1 Calomiris and Mamaysky, 2020

How Natural Language Processing Will Improve Central Banks' Accountability and Policy

Calomiris and Mamaysky explore the challenges of central banks' accountability and the clarity of their strategic communications. They argue that despite the amount of information provided by central banks, there is a lack of systematic framework in how textual data translates into actual policies. In particular, they claim that the Fed's reaction function to economic indicators has become more and more obscure over the past decades: FOMC minutes and Fed chair announcements have become increasingly technical and less intelligible. Paradoxically, the increasing volume of communication may have actually reduced transparency. This trend can be exemplified by former Fed Chairman Alan Greenspan's quote: "If you understood me, then I misspoke," reflecting a strategic choice to be vague.

They argue that NLP could significantly enhance the readability of central bank communications and strategies and thereby make market reactions more efficient. NLP methods could systematically analyze the speeches of central bankers to decode their implicit messages and anticipate their actions. These tools could be used to review the vast and increasing volume of textual data from central bank communications and identify subtle cues or specific patterns that indicate policy moves. They would therefore provide accurate forecasts for monetary policy, the same way forward guidance does.

For instance, they suggest using the so-called Prattle's Index, which places central bank sentiment, or estimated policy stance, on a spectrum ranging from "hawkish" to "dovish." According to them, this index has demonstrated its effectiveness by linking nuances of central bank sayings with movements in financial and economic indicators.

They highlight the prospective benefits of NLP for central bankers themselves, helping them save money and make economic forecasting and policymaking more efficient and quicker. Indeed, current NLP tools have demonstrated the ability to predict significant variances in macroeconomic indicators, such as GDP growth forecasts, up to 18 months in advance using only news data. This implies that an analysis of public information only could yield valuable insights, which could be particularly beneficial for central bankers who generally spend a lot of time and money to gather private data.

To conclude, leveraging NLP would help central banks bridge the gap between their exclusive data and the large amount of publicly available information, making policymaking more methodical and transparent. Moreover, this transformation of monetary policy formulation would make policies more predictable for market participants.

A.1.2 Hansen, McMahon, Prat, 2018

Transparency And Deliberation Within the FOMC: A Computational Linguistics Approach

Hansen, MacMahon, and Prat explore the effect of the introduction of transparency to the Fed's monetary policymakers' deliberations. They use a natural experiment from 1993, when FOMC transcripts became publicly available, and computational linguistics to assess two opposing effects of increased transparency: (i) the discipline effect, promoting more diligent work from policymakers (the positive effect), and (ii) the conformity effect, encouraging conformity, potentially stifling individual expression and debate (the negative effect).

Using machine learning algorithms, they found that FOMC's communication patterns notably changed post-transparency, identifying a shift towards more disciplined and data-driven discussions. They conducted a difference-in-differences analysis and identified that less experienced members exhibited both effects: a discipline effect in their prepared remarks (FOMC1) and a conformity effect during the policy strategy discussion (FOMC2).

Moreover, they constructed an influence measure inspired by the PageRank algorithm and showed that post-transparency, less experienced members' topics of interest influenced their colleagues more, suggesting a net positive/discipline impact of transparency on FOMC deliberations.

To conclude, they argue that transparency introduced some conformity and its overall effect seemed to have enhanced the deliberative process, particularly for less experienced members, leading to richer discussions. However, they nuance the conclusion, underlying that transparency's discipline effect may have been as or more significant than the conformity effect in some contexts.

A.1.3 Gáti, Handlan, 2023
Monetary Communication Rules

Gáti and Handlan explore the systematic "monetary communication rules" used by the Fed in its FOMC statements, focusing on how these communications correlate with the Fed's internal economic forecasts like inflation and GDP growth. Moving away from traditional narrative or dictionary-based text analysis, they adopt a quantitative text analysis framework employing a bag-of-words model and ridge regression. This approach allows them to link specific words and phrases in FOMC statements directly to changes in economic forecasts, revealing the systematic nature of the Fed's communications.

Their analysis highlights that different words and phrases are used deliberately to signal changes in economic policy and forecasts. The paper identifies significant shifts in the language of FOMC statements, particularly during periods of major economic events, such as the financial crisis in 2008. These shifts are shown to correlate with changes in market expectations and monetary policy surprises, indicating that changes in communication styles can significantly impact market reactions.

Furthermore, Gáti and Handlan suggest that the Fed's communication rules are not only systematic but also adaptable, with changes in these rules often aligned with shifts in the economic environment. They underscore the importance of understanding these communication rules to effectively interpret and predict the Fed's policy actions.

A.2 Explanatory Data Analysis

The data below reports the correlation (Figure 1) and multicollinearity using VIF (table 3) for all the selected features as potential candidates for the regression to estimate $\widehat{\mathbb{E}}_{t-\Delta t}[HD_t]$ in Equation 10.

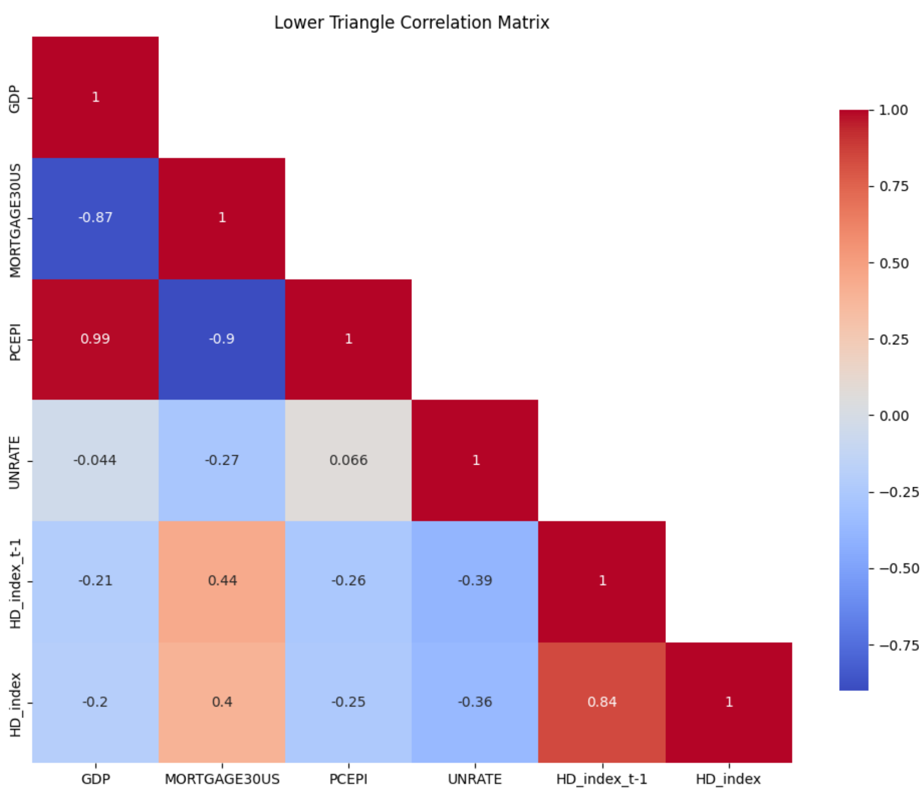

Figure 1: Correlation Matrix of Features

Table 4 below reports the multicollinearity of the selected features for the final regression to estimate $\widehat{\mathbb{E}}_{t-\Delta t}[HD_t]$ in Equation (10). We see VIF figures "under control".

Table 3: Variance Inflation Factor (VIF) for Selected Features.

Feature	**VIF**
GDP	706
MORTGAGE30US	87
PCEPI	1512
UNRATE	33
HD_{t-1}	1

Table 4: Variance Inflation Factor (VIF) for Features of Final Model

Feature	**VIF**
GDP_GROWTH	1.2
UNRATE	1.2
HD_{t-1}	1.0

A.3 Results

A.3.1 Number of Dimensions of Markets' Response to Monetary Policy Announcements

Table 5: Summary of Rank Test Results for All Datasets

Dataset	**Rank**	**Min Distance**	**Deg of Freedom**	**Critical Value**	**p-Value**
Dataset 1 Bonds, Forex (7 columns)	0	93.60	21	32.67	0.00
	1	40.39	14	23.68	0.00
	2	**13.93**	**8**	**15.51**	**0.08**
	3	5.90	3	7.81	0.12
Dataset 2 Bonds, Forex, S&P500 (8 columns)	0	98.23	28	41.34	0.00
	1	53.09	20	31.41	0.00
	2	**22.40**	**13**	**22.36**	**0.05**
	3	8.24	7	14.07	0.31
Dataset 3 Bonds, Forex, NASDAQ (8 columns)	0	107.59	28	41.34	0.00
	1	54.87	20	31.41	0.00
	2	**16.94**	**13**	**22.36**	**0.20**
	3	8.17	7	14.07	0.32
Full Dataset Bonds, Forex, Equities (9 columns)	0	121.89	36	51.00	0.00
	1	59.60	27	40.11	0.00
	2	**29.83**	**19**	**30.14**	**0.05**
	3	16.20	12	21.03	0.18

A.3.2 Independent Component Analysis

In Figures 2 and 3, the y-axis displays the features, and the x-axis denotes the components; Independent Component 1 (IC1) corresponds to the left column and Independent Component 2 (IC2) corresponds to the right column. Features with 'gt' indicate treasury bonds on a yearly level, with the maturity following right after. Features with'gb' indicate bonds on a monthly level, with its maturity denoted right after. The number '2' and last are just preprocessing indications and hold no significant interpretation. E.g. 'gt_10_2_last' denotes the 10-year treasury bond feature, and gb_06_2_last denotes the 6 month bond feature. 'eurusd' denotes the currency exchange between euros and dollars. Lastly, 'tot_spx' and 'tech_ndx' indicate the S&P500 index, and a combination of tech companies from the NASDAQ index respectively.

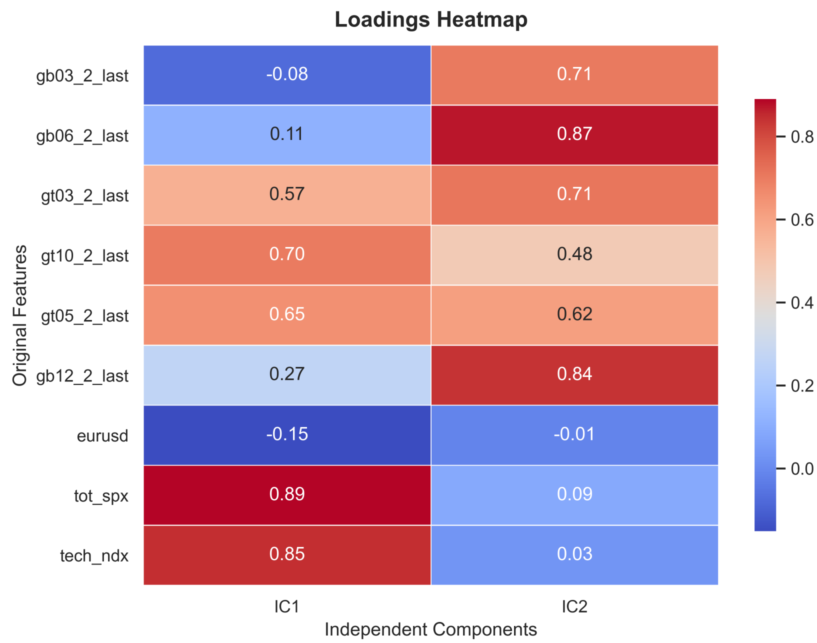

Figure 2: Loadings of Components on Features

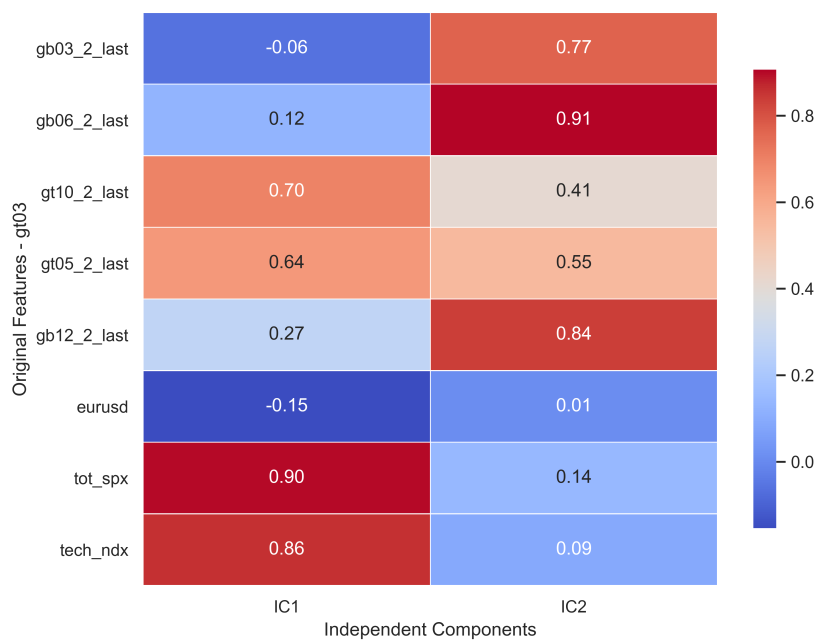

Figure 3: Stability Test, Loadings of Components on Features without the 3 year maturity bond

Figure 3 displays the result of a stability test of the ICE, with loadings of components on features without the three year maturity note.

A.3.3 Structural Interpretation of the Two Factors

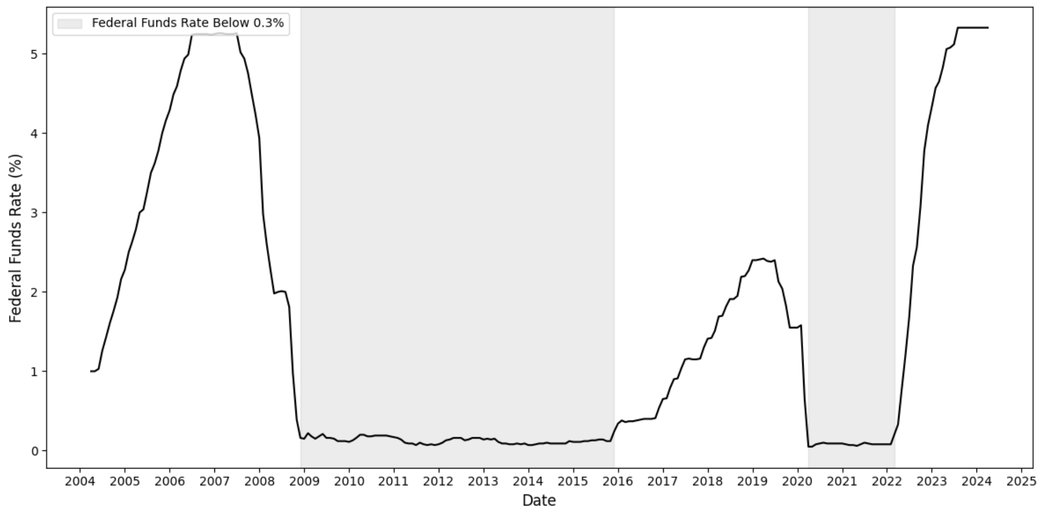

Figure 4: Federal Funds Effective Rate (2004 - 2024)

A.3.4 A New Measure for Market Surprise

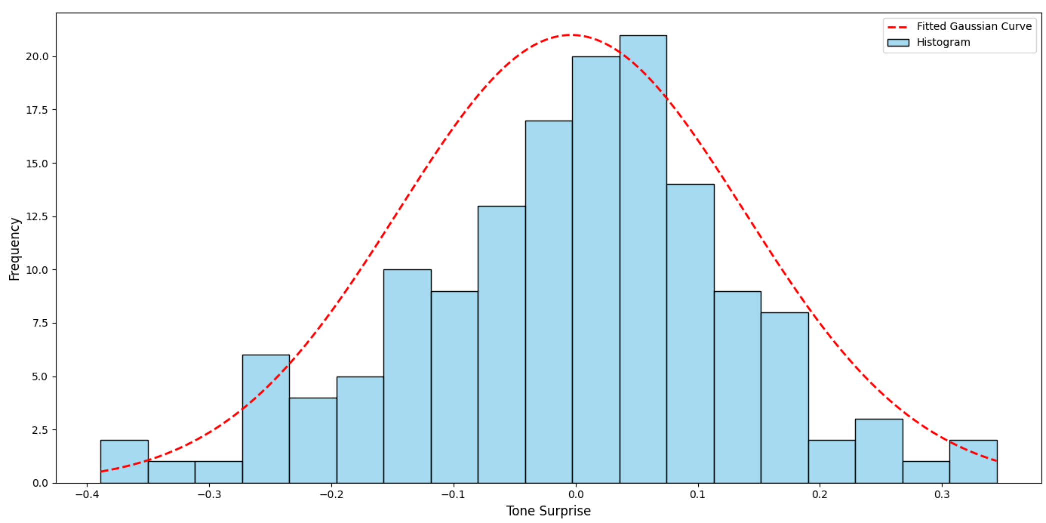

Figure 5: Frequency Distribution of Tone Surprise with Fitted Gaussian Curve

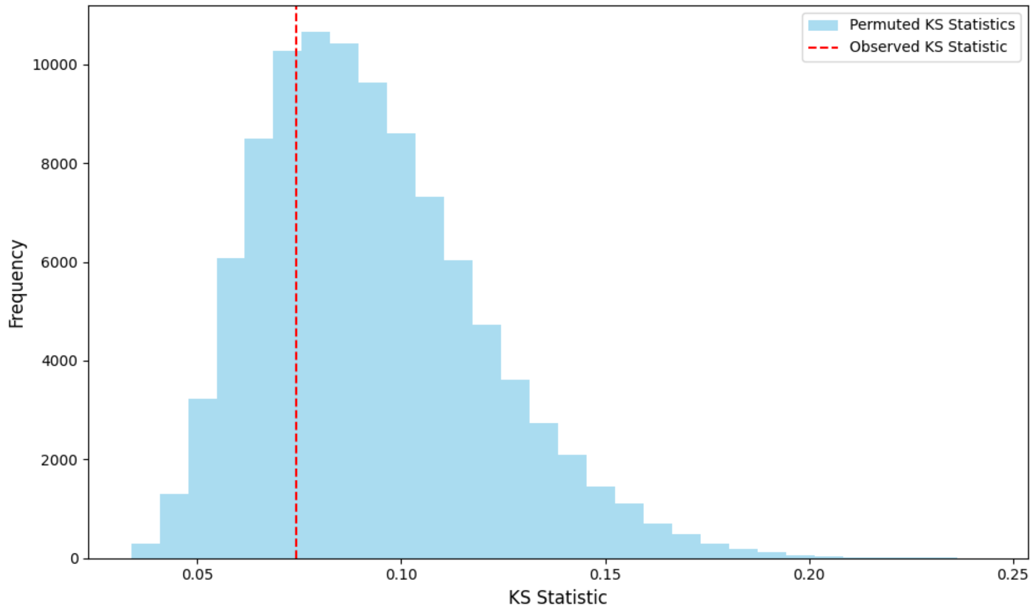

Figure 6: Results of the Permutation test with the KS Statistic

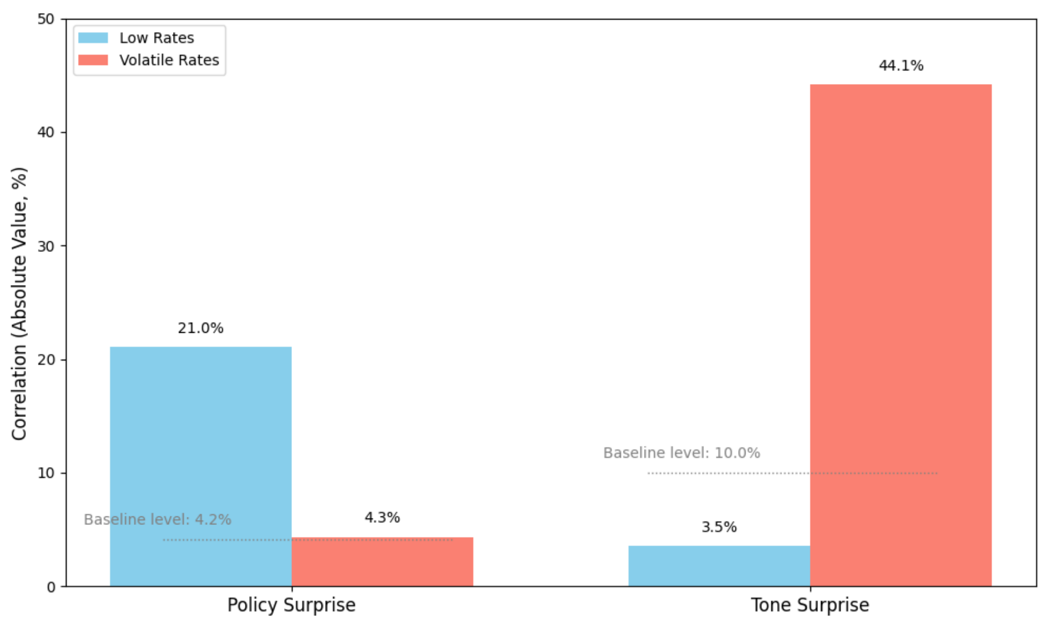

Figure 7: Correlation between the first column of the adequately rotated matrix and market surprise measures by rate environment