Reading Between the Lines

Using Sentiment Analysis to Understand Future Federal Funds Rate Changes.

Carl Somers

19422704

Supervised by Dr. Oana Peia

Abstract

This study investigates the information value of Federal Open Market Committee (FOMC) meeting minutes and their potential to provide systematic indications of future policy decisions. I apply Natural Language Processing techniques to the meeting minutes to identify information that is not apparent through traditional quantitative measures. I then estimate a Taylor-Rule model that incorporates the sentiment of the meetings to analyse Fed Funds Rate changes between March 1996 to February 2023. My findings suggest that this information can explain future interest rate changes beyond what is explained by traditional macroeconomic variables.

Keywords: FOMC, Monetary Policy, Sentiment Analysis, Federal Funds Rate, Natural Language Processing, Taylor Rule

I. Introduction

Central banks across the world have increasingly relied on public communications to provide guidance on future policy actions (Casiraghi & Perez, 2022). Where market participants once had to infer beliefs around the path of policy using quantitative information, such as interest rates or inflation, there is now access to rich qualitative data. The practice of open communication became even more critical with the emergence of the effective zero-lower bound on interest rates, beginning in the US in 2007 (Bernanke, 2010). As a result, the Federal Reserve (Fed) lies at the forefront of this trend.

Nowadays, the Fed is ranked among the most transparent central banks in the world. The Federal Open Market Committee (FOMC) is the committee responsible for overseeing open market operations and meets eight times a year to make decisions on the target federal funds rate and money supply. Along with a statement presenting a brief on the decisions, they maintain a record of meeting minutes, released 3-weeks later. These minutes contain a more detailed account of the policy makers economic outlook and the reasoning behind their verdicts. The understanding of this qualitative information has become an area of critical importance, however, extracting meaningful insights remains a challenging task given their length and complexity.

In recent years, the advancements of natural language processing (NLP) techniques have reshaped research in this information rich environment. Compared to a manual approach such as Romer and Romer (1989), technology can help minimise potential researcher bias. Using historical FOMC meeting minutes, my paper outlines the use of automated sentiment analysis to extract essential information, converting this qualitative information into quantitative metrics. This analysis, sometimes referred to as sentiment analysis, seeks to identify the emotional tone behind a text communication.

For my empirical analysis, I explore over 27-years of text data from the FOMC minutes using sentiment analyses. This text data was from the FOMC website and collected using a custom python web-scraping system. Not only do I use a bag-of-words approach and the financial dictionary proposed by Loughran and Mcdonald (2011) (LM) to measure the positive/negative sentiment, but I also employ a more advanced FinBERT model (Araci, 2019). FinBERT is a pre-trained financial language model which has demonstrated context specific sentiment analysis. Currently NLP models mostly focus on word-level or sentence-level analysis, therefore, I explore a novel, document level approach.

I chose to use my sentiment measure to understand future changes to the Federal Funds Rate (FFR) due to its unquestionable importance to both the US and the global economy. The FFR is often referred to as the "policy rate" because it is the primary tool used by the Federal Reserve to implement monetary policy and influence economic conditions. Increasing interest rates can help cool an overheating economy, while reduction can be stimulating. Adjustments to the FFR are made to achieve the central bank's dual mandate of maintaining price stability and maximizing employment.

Therefore, to anticipate the appropriate level of interest rates, I employ the widely used Taylor rule, a popular framework in monetary policy analysis that considers the current levels of inflation and output gap (Taylor, 1993). My model estimates the probabilities of a change in FFR while accounting for these conventional macroeconomic variables. By employing such variable controls, my model provides a nuanced understanding of the prevailing economic environment during each policy meeting, which is critical for comprehending the Federal Reserve's monetary policy decisions.

My empirical research presents the importance of analysing FOMC minutes, especially their sentiment, to garner expectations for future interest rate changes. I found that sentiment from these minutes has statistical significance in the understanding of these future changes. This prediction is robust to different measurement of sentiment, period resampling and even the inclusion of macroeconomic variables. This has several crucial implications from the uses in macroeconomic forecasting, to the importance of FOMC communication for understanding future policy.

My study extends the existing literature by conducting a comprehensive set of robustness checks to validate my findings. In particular, I introduce a novel approach by incorporating the FinBERT model to calculate sentiment scores. I also use more robust samples of data, controlling for the Zero Lower-Bound on interest rates and including data up to 2023. My research employs data science techniques in the understanding of economic data.

The remainder of the paper is organised as follows. Section 2 reviews the body of research relating to three related areas: text mining for central bank communication, the effectiveness of central bank communication and finally macroeconomic forecasting. In Section 3, I explain why the minutes may contain unique information in comparison with other forms of FOMC communication. In Section 4, my sentiment measure is described in more detail. Section 5 offers econometric evidence on the information content of the minutes along with key robustness checks, and Section 6 concludes reflecting on the implications.

II. Related Literature

My paper makes significant contributions to three strands of economic literature. First, it builds upon the growing trend of applying data mining techniques to central bank communications, particularly sentiment analysis, to uncover policymakers' decisions. Researchers have used these methods on various forms of communication, from text-based to facial recognition (see Curti and Kazinnik, 2022), to quantitatively capture traditionally qualitative information.

Shapiro and Wilson (2021) propose a new approach to estimate the objectives of central banks, specifically the Fed. By using Latent Dirichlet Allocation (LDA) they identify and classify topics from FOMC statements. They observe how these objectives have changed over time, as the behaviour and challenges faced by the Fed have evolved.

Gardner, Scotti and Vega (2022) analyse how financial markets react to the description of the economy painted by the FOMC statements. Using their sentiment measure they note that during bad times, the macroeconomic outlook resulted in a larger change in equity prices than during good times. They also suggest that this effect is related to the use of the FOMC's economic outlook as a predictor for interest rate changes.

Second, my paper contributes to the literature on the effectiveness of central bank communication as a monetary policy instrument, specifically its ability to provide forward guidance. Policymakers use forward guidance to signal their future intentions for monetary policy, but its effectiveness depends on the clarity and accuracy of the information provided to the public.

Blinder et al. (2008) argue that central bank communication can be an effective monetary policy tool if it is clear and consistent. Their research points towards markets digestion of central bank communication through enhanced predictability of monetary policy and its potential to achieve their macroeconomic objectives. This highlights the importance of using language that is easily understood by the public and the need for transparency and accountability in communication. However, they also note that the consensus on the optimal communication has not yet emerged.

Jegadeesh and Wu (2017) employ a LDA model to extract common themes from FOMC minutes. They find that text mining from the minutes provides relative informativeness about the path of FOMC policy through the reactions of the stock market. Their results suggest that text mining techniques can enhance the accuracy and aids the transmission of forward guidance provided by central banks. These techniques can therefore increase public understanding of both past and future FOMC decisions and aid in meeting Fed mandates.

Other studies have explored the effectiveness of central bank communication in different contexts. Jung (2016) analyses the impact of ECB communications, finding that they are an effective tool in satisfying information demands and anchoring expectations. Gürkaynak, Sack and Swanson (2018) investigate the effects of central bank communication on financial market expectations, finding that it can influence market perceptions of future monetary policy.

Finally, my paper contributes to macroeconomic forecasting, specifically the Fed Funds Rate. Taylor's (1993) simple monetary policy rule based on the fed funds rate has become widely used as a benchmark for monetary policy analysis. His research shows that changes in the fed funds rate are strongly correlated with changes in inflation and output, and that a simple rule-based approach can achieve good economic outcomes. I employ this benchmark to discern predictive information of Fed Funds Rate greater than traditional methods in my paper.

The use of sentiment analysis on central bank communication for predicting macroeconomic variables, especially interest rates, is a growing field. Hubert et al. (2021) observes the tone of FOMC statements and find it useful for predicting future policy decisions. They also note that sentiment analysis explains monetary surprises, such as interest rates, beyond those seen in policymakers' forecasts and voting records. Additionally, they find that sentiment is a particularly informative variable for forecasting at monetary cycle turning points.

Combining all three associated fields, my work is closely related to Apel and Grimaldi (2012). These researchers employ a bag-of-words method to extract a *Net Index* sentiment measure from Central Bank of Sweden (Riksbank) minutes. They found that Riksbank minutes contain valuable information for predicting interest rate changes, even when accounting for basic macroeconomic variables and market participants' expectations. However, I improve upon the shortcomings of their paper through my robustness tests employing a more novel modelling approach (FinBERT) and Zero-Lower Bound interest rate period exclusions.

III. Background on FOMC Meeting Minutes

The 1990's saw a rise in congressional interest in the FOMC information distribution leading to the creation of the meeting minutes in 1993. At this time, research on the importance of public expectations on policy implementation was at the forefront of economic literature, championed by the likes of Milton Friedman (see Friedman, 1968). These minutes aimed to increase public understanding of monetary policy and its future path.

In 2004, the FOMC halved the time between the meeting and publication of their meeting minutes from 6-weeks to 3-weeks, this heightened public attention to the minutes, and increased their possible informativeness. Committee members noted not only do these minutes hold a more complete description of the FOMC's outlook, but also that an earlier release would allow markets to interpret economic developments and predict the course of interest rates (Danker and Luecke, 2005). It is with these elements in mind that I employ text mining for the extraction of information from these minutes.

Note, these are not the sole form of public information, and I am conscious of the shortcomings of not analysing statements, transcripts, speeches and the vast array of FOMC communication as well. Minutes have significant advantages such as text length, and timely release which allow them to be used for the extraction of more robust sentiment analysis. However, these minutes are still aptly tailored summaries and therefore may not entirely capture the FOMC outlook.

IV. Quantifying Qualitative Information

In this section, I will first present the data sources used for subsequent analysis. Next, I will explain the construction of a sentiment score for FOMC minutes.

4.1 Text Data

In this study, I obtained the FOMC meeting minutes spanning from March 1996 to February 2023, which comprise a total of 213 documents. To collect this data, I systematically downloaded the documents from the FOMC website and extracted the text using a custom web-scraper. This automated process enabled us to efficiently retrieve website data and export it in a structured format. Additionally, the date of release and the Fed Chairman who presided over each meeting was recorded. I will refer to this dataset as "Minutes" throughout the paper.

Table 1 presents the average word count across different chair mandates.

	Average Word Count	*Beginning*	*End*
Alan Greenspan	4939	1987-08-11	2006-01-31
Ben Bernanke	6786	2006-02-01	2014-01-31
Janet Yellen	8754	2014-02-03	2018-02-03
Jerome Powell	9263	2018-02-05	Present

Table 1 - Average Minute Word-Count by Fed Chair

It is clear to see that the minutes themselves have seen a large structural change through the various Fed chairpersons. This is a recognition of the increasingly transparent nature of central banking (Dincer and Eichengreen, 2013). However, much of this communication could be insubstantial which would introduce undue noise to my modelling.

For these Minutes, an automated method to parse the noisy data was developed. This is a several step process as follows: 1) Break the document into paragraphs and count words; 2) Remove irrelevant sections below 50 words, such as the introductory section which lists participants names and outlines administrative issues; 3) Clean remaining text by removing words providing no additional information referred to as stop-words (e.g. “the”, “because”, “and”), remove punctuation, and tokenise each individual word.

To demonstrate the power of this basic text cleaning, the word-counts of the raw and the cleaned text can be observed in table 2. See Appendix A for an example of this text cleaning.

	Raw FOMC Minutes	*Clean FOMC Minutes*
Mean Word Count	6843.74	3387.55

Table 2 – Mean word count pre- and post-cleaning.

4.2 Automated Content Analysis

To quantify the tone of the FOMC Minutes, I use an automated content analysis approach known as the bag-of-words method. To implement the bag-of-words approach, this research relies on a dictionary of predefined words created by Loughran and Mcdonald (2011) (LM), which has been extensively used across the literature (e.g Apel and Grimaldi, 2012; Tadle, 2022; Bennani and Romelli, 2021). This dictionary classifies words tone as either positive or negative developed using corporate 10-K text reports between 1994 and 2008 and the proceeding returns and volatility.

In line with previous literature, the net sentiment of the Minutes is classifed as either hawkish or dovish. Hawkish sentiment indicates an optimistic outlook on the economy, characterized by improved growth and rising inflation, which implies a monetary tightening. On the other hand, dovish sentiment suggests a bleaker economic outlook, which *typically* calls for a loosening of monetary policy.

To account for instances where a leading negation reverses the sentiment of a keyword, I incorporate a negate specification. For example, the phrase "not improving" is recognized as dovish, even though the keyword "improving" is typically associated with a hawkish sentiment.

The hawkish (dovish) sentiment of each meeting is computed by measuring the frequency of positive (negative) keywords across all the Minutes.

Figure 1 - Word-Cloud of Hawkish Terms

Figure 2 - Word-Cloud of Dovish Terms

Finally, to measure the intensity of these dovish/hawkish terms used in the Minutes according to the LM dictionary, a net sentiment measurement is calculated. This meeting *sentiment* can be quantified as follows (Bennani and Romelli, 2021):

$$Sentiment_m = \frac{hawk_m - dove_m}{hawk_m + dove_m}$$

Where *m* refers to the current meeting minutes, *hawk* refers to the document's count of hawkish classified words, and *dove* corresponds to the dovish classification. This net sentiment calculation weighs the ratio of hawkish and dovish sentiment across each document in my Minutes dataset. Therefore, overall outlook is measured on a continuous basis. To ensure this continuous measure is scale free, it is transformed using standard scaling. This is a common technique that achieves generalised statistical properties with a mean of zero and a standard deviation of 1.

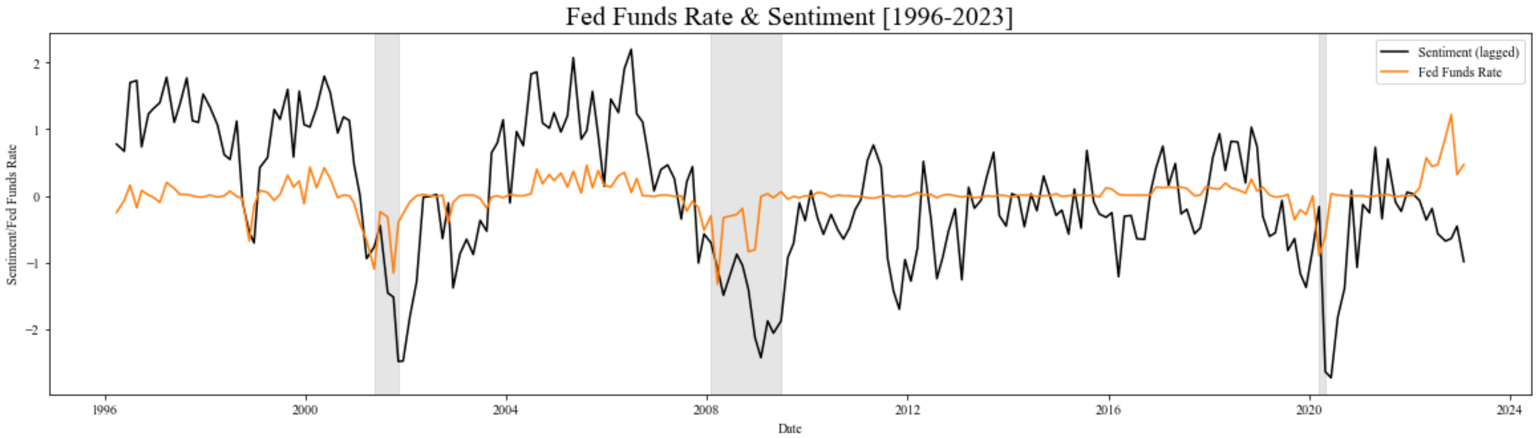

Figure 3 - Net Sentiment and FFR changes, NBER Recessions highlighted.

As shown in figure 3, there is a clear correlation between the lagged net sentiment of the FOMC and the change in interest rate. The sentiment is lagged to correspond to the previous meeting. My analysis reveals that dovish sentiment is *generally* associated with policy rate decreases and vice versa, consistent with historical patterns. In periods of recession, denoted by the grey highlights, sentiment is largely negative, and the interest rates enter a recessionary regime.

V. Empirical Evidence

At each meeting the FOMC faces 3 distinct possibilities: to tighten monetary policy, to loosen or to keep it unchanged. As changes to the fed funds rate are in multiples of 25 basis points (bps or 0.0x%) and changes greater than this are rare (Poole, 2005), the variables for Δr_{t+k} can be transformed into ternary/discrete variables. Therefore:

$$\Delta r_{t+k} = \begin{cases} +25bps & \Delta r_{t+k} > 0.025\% \\ 0 & -0.025\% < \Delta r_{t+k} < 0.025\% \\ -25bps & \Delta r_{t+k} < -0.025\% \end{cases}$$

The ordered probit is used commonly in literature with discretionary dependent variables (Jung, 2016; El-Shagi & Jung, 2015; Horváth et al., 2012). Therefore, to investigate the predictive contents of the minutes for the path of future policy decisions, I will evaluate the model with this simple specification (Apel and Grimaldi, 2012):

$$\Delta r_{t+1} = \beta_1 \Delta r_t + \beta_2 \text{Sentiment}_t + \varepsilon_t$$

Δr_{t+1} is the interest rate policy decision at time *t* + 1, the next policy meeting, Δr_t is the most recent policy decision, *sentiment* is the net tone score for the most recent FOMC minutes (time *t*) and ε_t is the stochastic error term.

The changes in future policy decision are regressed on the sentiment variable while controlling for the previous policy change. The results are presented in table 3.

Ordered Probit Results

	Dependent variable: $\Delta r_t + 1$		
	Naive Model	Sentiment	Hawkish/Dovish
	(1)	(2)	(3)
Δr_t	0.058***	0.045***	0.047***
	(0.009)	(0.009)	(0.009)
$Sentiment_t$		0.466***	
		(0.117)	
$Hawkish_t$			0.006**
			(0.003)
$Dovish_t$			-0.007***
			(0.002)
-25bps	-1.566***	-1.729***	-1.769***
	(0.143)	(0.166)	(0.464)
+25bps	1.119***	1.215***	1.191***
	(0.066)	(0.073)	(0.071)
Observations	214	214	214
Pseudo R^2	0.18	0.24	0.23
AIC	224.9	209.5	215.7
Log-Likelihood	-109.46	-100.76	-102.84
Note:		*p<0.1; **p<0.05; ***p<0.01	

Table 3 - Ordered Probit Estimation including Sentiment.

From my estimation the relationship between FOMC minutes sentiment and the future rate path of interest rate policy is observed. Note that lagged interest rate changes and sentiment measures – both net and separated (dovish/hawkish) are statistically significant at all levels (α = 1%, 5%, 10%) in the prediction of future rate changes. Therefore, the sentiment of the FOMC is an extremely useful indicator for the path of policy. This is further supported by the incremental improvements in reported pseudo-R^2, log-likelihood and AIC with the inclusion of sentiment variables in the models[1].

The cut-off point estimates (+25bps & -25bps) are both significant, these represent the switching point of the ordering. The intercepts here show the probability of the rate tightening, loosening, or remaining neutral while all explanatory variables are zero. The larger spread between the cut-off points, indicate a stronger impact of the sentiment variables on the probabilities.

The ordered probit provides some beneficial statistical properties that make it robust for modelling, however, as the coefficients do not correspond to marginal contributions seen in an ordinary least squares (OLS) regression (Wooldridge, 2012)[2]. These are instead interpreted as changes to the log-odds based on one unit change of the predictor variable, while holding other variables constant.

For example, using this simple specification the probability of a rate decrease is calculated using the cut-off/ancillary variables (Apel and Grimaldi, 2012; Wooldridge, 2012):

$$\mathbb{P}(\Delta r_{t+1} = -0.025\%) \ = \ \Phi(-25bps \ - \ \beta_2 \text{Sentiment}_t \ - \ \beta_1 \Delta r_t)$$

Here Φ denotes the normal cumulative density function. If the sentiment measure is held at its mean level and the previous rate decision is unchanged, the probability of a rate reduction is 4%. While the probability of an increase in interest rates is 11%, given by:

$$\mathbb{P}(\Delta r_{t+1} = 0.025\%) \ = \ 1 \ - \ \Phi(+25bps \ - \ \beta_2 \text{Sentiment}_t \ - \ \beta_1 \Delta r_t)$$

---

[1] $Pseudo - R^2 = 1 - \frac{Log\ Likelihood_{model}}{Log\ Likelihood_{null}}$
This is the McFadden R^2 and is a goodness of fit metric for ordinal models. It is suggested that a $Pseudo - R^2$ between 0.2 and 0.4 indicates an extremely good fit. McFadden and Domenicich (1975) find through simulations this range can be compared to a 0.7 to 0.9 for a linear functions R^2.

[2] Specifically, these statistical benefits/properties include capturing non-linear relationships, ability to handle ordinal dependent variables and accounting for unobserved heterogeneity.

Finally, no change in interest rate is 85% and calculated as:

$$\mathbb{P}(\Delta r_{t+1} = 0) = \Phi(+25bps\ -\ \beta_2 \text{Sentiment}_t - \beta_1 \Delta r_t)\ -\ \Phi(-25bps\ -\ \beta_2 \text{Sentiment}_t \\ -\ \beta_1 \Delta r_t)$$

The probabilities obtained from the probit model align with historical trends of static interest rates in no-change regimes which is reflected in the high proportion of no-change outcomes in the sample (Audrino and Offner, 2021). However, when the sentiment shifts to hawkish, one standard deviation above the mean, the probabilities skew towards a 23% chance of a rate increase, a 1% chance of a rate decrease, and a 76% chance of returning to neutral. Conversely, a dovish sentiment results in a 10% chance of a rate decrease, a 6% chance of a rate increase, and an 84% chance of remaining at neutral.

These results suggest that incorporating sentiment from the FOMC minutes into the analysis strongly improves the prediction of future interest rate changes.

VI. Robustness Tests

In addition to the main empirical exercise, I also consider three robustness checks. First, a model employing macroeconomic information into the original specification is estimated. Second, different calculations for sentiment are modelled. Finally, the zero lower bound (ZLB) period is removed from the sample.

6.1 Macroeconomic Specifications

For my sample dates March 1996 to February 2023, I download the corresponding macroeconomic data from FRED. I retrieve my dependent variable, the Federal funds rate (FFR), along with the Consumer Price Index (CPI), US Gross Domestic Product (GDP), and potential GDP. It is important to note that the Federal funds rate is measured on a 6-weekly timeline, while inflation is reported monthly and GDP on a quarterly basis.

GDP and potential GDP are used to calculate the output-gap. This is a more informative measure as it accounts for the potential of the economy given its available resources such as labour, technology and capital. This represents the difference between the current economy and the maximum sustainable level of output without generating inflationary pressure.

The summary of this data can be observed in Table 4 below.

	Mean	*Standard Dev.*	*Min*	*Max*
% Δ Inflation	0.21	0.34	-1.77	1.96
Output Gap	0.75	2.48	-5.68	10.82
% Δ Fed Funds	0	0.27	-1.33	1.22

Table 4 – Macroeconomic data summary statistics

As my sentiment measure is intended to predict future policy changes, it is essential to determine whether it provides any additional information beyond that captured by traditional macroeconomic variables. Specifically, I will test whether the relationship between my sentiment measure and the Federal funds rate remains significant after accounting for the effects of other explanatory variables. To do so, I will augment my baseline model with a Taylor-type rule that includes additional macroeconomic variables such as inflation and the output gap. By comparing the results of this augmented model to the baseline, I can determine whether my sentiment measure holds any excess information beyond that already captured by macroeconomic data. This Taylor-type model is specified as:

$$\Delta r_{t+1} = \beta_1 \Delta r_t + \beta_2 \text{Sentiment}_t + \beta_3 \text{Inflation}_t + \beta_4 \text{OutputGap}_t + \varepsilon_t$$

Here I have added the inflation rate and output gap at time *t* to my initial specification. I re-sampled both to a 6-week frequency due to time sampling differences - GDP is a reported quarterly and CPI monthly. In doing this I hope to employ the data as it will be used in my prediction model, that is the last known data at time *t* with no lookahead bias.

Ordered Probit Results

	Dependent variable: $\Delta r_t + 1$			
	Naive Model	Sentiment	Taylor Rule	Sentiment-Taylor Rule
	(1)	(2)	(3)	(4)
Δr_t	0.058***	0.045***	0.056***	0.043***
	(0.009)	(0.009)	(0.009)	(0.010)
$Sentiment_t$		0.466***		0.610***
		(0.117)		(0.143)
$Inflation_t$			0.691**	0.588*
			(0.339)	(0.352)
$OutputGap_t$			-0.000	0.144**
			(0.048)	(0.061)
-25bps	-1.566***	-1.729***	-1.449***	-1.474***
	(0.143)	(0.166)	(0.169)	(0.190)
+25bps	1.119***	1.215***	1.136***	1.242***
	(0.066)	(0.073)	(0.067)	(0.075)
Observations	214	214	213	213
Pseudo R^2	0.18	0.24	0.20	0.27
AIC	224.9	209.5	223.3	205.3
Log-Likelihood	-109.46	-100.76	-106.66	-96.66
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5 – Ordered Probit with Macroeconomic Specification

From the modified model, the sentiment measure is still significant at all levels (α = 1%, 5%, 10%) when accounting for macroeconomic information. This shows that even when accounting for the basic macroeconomic outlook minutes sentiments remain informative. Comparing the pseudo-R^2 estimates to the core sentiment model confirms the fitting improvement from adding macroeconomic data.

6.2 Alternative Sentiment Measurements

Although dictionary-based indicators have been widely used in the financial and economic domain, there are concerns about the reliability of this traditional method. Largely this method reduces each word to a single dimension, falling short of analysing deeper semantic meaning of a text. According to Hayo and Zahner (2023), this approach is flawed and can attribute up to 80% of measurement variation to noise. Therefore, to enhance the robustness of my sentiment analysis, I will estimate models using FinBERT, a state-of-the-art deep learning model designed specifically for financial text analysis.

FinBERT is a pretrained model used to analyse sentiment of financial text. This model is a variation of Google's BERT (Bidirectional Encoder Representations from Transformers), a ground-breaking development to language modelling (Chang et Al., 2019). This model uses deep neural network architecture to learn the relationship between words and their context, the revolutionary introduction of *bidirection* means it can look at surrounding words and infer context both directions (left and right). FinBERT has been trained on a specific large corpus of financial text and is therefore tuned for sentiment in this domain (Araci, 2019). As finBERT has a 528 token maximum, overlapping text batches are used to calculate sentiment for the whole document without losing context. This raw model is applied without topic specific fine-tuning on the FOMC minutes.

The results from this model are observed in the Appendix B, Table 6.

Using the raw finBERT model to calculate sentiment, the estimated model underperforms in comparison to the simple bag-of-words method. This is demonstrated by the reduction in pseudo-R^2 and increase in AIC. From my coefficients I observe a lower impact of FOMC sentiment on the selection of future interest rates. However, the measure remains statistically significant, this ultimately adds validity to the quantitative measurement of sentiment from the FOMC minutes and suggests that it is a robust measurement for prediction of future policy changes.

6.3 Shifts in Textual Sample over Time

As discussed previously, an overrepresentation of neutral/no rate change policy could result in a skewed prediction model. The zero-lower bound (ZLB) on interest rates plays a crucial role in this data imbalance within the sample period.

The ZLB is a monetary policy tool used by central banks to lower short-term interest rates, such as the Fed funds rate, to zero in an effort to stimulate economic growth. While this policy was in effect in the United States between 2008 and 2015, other central banks, including the Bank of Sweden, Japan, and the European Central Bank, implemented negative interest rates in an attempt to spur economic recovery.

As interest rates are a critical tool in central banking policy, when the ZLB rate is ineffective in improving economic conditions, the central bank may resort to unconventional measures, such as quantitative easing. It is worth noting that during the ZLB period, changes in sentiment may not be accurately captured by the Fed funds rate. Therefore, these periods are excluded, with a new sample size of 141.

The results from this model are observed in the Appendix B, Table 7.

From the modelling of the refined sample, there is a continued robustness of my sentiment variables on future interest rate changes. Once again, there is statistical significance for all cut-off and coefficients at all levels. The ancillary/cut-off points suggest a lesser effect of sentiment on the selection of ordinal category. Therefore, it is no surprise this model presents a diminished pseudo-R^2 measure indicating worse model fit with the new sample. This could be a result of the increased variance within the interest rate changes once the ZLB is removed. Observing the AIC, however, note there is an improvement compared to models accounting for the entire sample period, this is likely caused by the reduced sample size.

VII. Conclusion

Central bank communications have increased their relevance in recent policy making. These documents aid in the communication of the path of future policy decisions, outlining the economic outlook and informing on the decision-making process. One of the key communication developments is the release of the FOMC meeting minutes. This empirical research presents the importance of analysing these minutes, especially their sentiment, to garner expectations for future interest rate changes. I found that sentiment from these minutes has statistical significance in the prediction of these changes.

These findings have critical implications for the importance of policy meeting minutes in modern central banking and transparent communication. Several major central banking institutions still do not provide these minutes, this includes the Bank of Japan, Swiss National Bank and Central Bank of United Arab Emirates. Through my analysis I show that institutions who release these minutes have increased openness, allowing for enhanced predictability of monetary policy.

I conducted multiple robustness tests to validate my results and found that my sentiment measure exhibited strong predictive power in forecasting changes to the Federal funds rate. Specifically, the researched measure outperformed traditional macroeconomic variables, remained robust despite data imbalances caused by the ZLB on interest rates, and across different sentiment measurements. These findings demonstrate the robustness of sentiment analysis and its potential as a valuable tool for economic forecasting.

To date, academic research regarding central bank communications has been limited to rudimentary approaches such as the bag-of-words, search & count or LDA methods. This paper demonstrates the use of a more contemporary FinBERT model for the semantic understanding of central banking texts. This opens the door to potential further research on the development of a fine-tuned model for this application. Uncovering the potential of my model to understand future interest rates, more advanced machine learning models can be employed to develop prediction models with a wider range of variables.

Bibliography

Apel, M. and Grimaldi, M. (2012) 'The Information Content of Central Bank Minutes', *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.2092575.

Araci, D. (2019) 'FinBERT: Financial Sentiment Analysis with Pre-trained Language Models'. arXiv. Available at: http://arxiv.org/abs/1908.10063 (Accessed: 10 April 2023).

Audrino, F. and Offner, E. (2021) *The impact of macroeconomic news sentiment on interest rates*.

Bennani, H. and Romelli, D. (2021) 'Disagreement inside the FOMC: New Insights from Tone Analysis'.

Bernanke, B. S. (2010, May). Central bank independence, transparency, and accountability. In Speech at the Institute for Monetary and Economic Studies International Conference, Bank of Japan, Tokyo, Japan, May (Vol. 25).

Casiraghi, M., & Perez, L. P. (2022). Central Bank Communications.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Curti, F. and Kazinnik, S. (2022) 'Let's Face It: Quantifying the Impact of Nonverbal Communication in FOMC Press Conferences'. Rochester, NY. Available at: https://doi.org/10.2139/ssrn.3782239.

Danker, D.J. and Luecke, M.M. (2005) 'Background on FOMC Meeting Minutes', *Federal Reserve Bulletin*, 91(2), pp. 0–0. Available at: https://doi.org/10.17016/bulletin.2005.91-2-2.

Dincer, N. and Eichengreen, B. (2013) 'Central Bank Transparency and Independence: Updates and New Measures', *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.2579544.

Friedman, M. (1968). The role of monetary policy (pp. 215-231). Macmillan Education UK.

Gardner, B., Scotti, C. and Vega, C. (2022) 'Words speak as loudly as actions: Central bank communication and the response of equity prices to macroeconomic announcements', *Journal of Econometrics*, 231(2), pp. 387–409. Available at: https://doi.org/10.1016/j.jeconom.2021.07.014.

Gürkaynak, R.S., Sack, B. and Swanson, E. (2018) 'Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements - IJCB - May 2005', *Premier issue (May 2005) of the International Journal of Central Banking* [Preprint]. Available at: https://www.ijcb.org/journal/ijcb05q2a2.htm.

Hayo, B., & Zahner, J. (2023). What is that noise? Analysing sentiment-based variation in central bank communication. Economics Letters, 222, 110962.

Jegadeesh, N. and Wu, D. (2017) 'Deciphering Fedspeak: The Information Content of FOMC Meetings', *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.2939937.

Jung, A. (2016) ‘Have FOMC Minutes Helped Markets to Predict FED Funds Rate Changes?’, *SSRN Electronic Journal* [Preprint]. Available at: https://doi.org/10.2139/ssrn.2839837.

Loughran, T. and Mcdonald, B. (2011) ‘When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks’, *The Journal of Finance*, 66(1), pp. 35–65. Available at: https://doi.org/10.1111/j.1540-6261.2010.01625.x.

McFadden, D. and Domenicich (1975) *Urban Travel Demand: A Behavioral Analysis, by Tom Domencich and Daniel McFadden, 1975, North-Holland*. Available at: https://eml.berkeley.edu/~mcfadden/travel.html (Accessed: 27 April 2023).

Romer, C.D. and Romer, D.H. (1989) ‘Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz’, in *NBER Macroeconomics Annual 1989, Volume 4*. MIT Press, pp. 121–184. Available at: https://doi.org/10.1086/654103.

Shapiro, A.H. and Wilson, D.J. (2021) ‘Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives Using Text Analysis’, *Federal Reserve Bank of San Francisco, Working Paper Series*, pp. 01–88. Available at: https://doi.org/10.24148/wp2019-02.

Tadle, R.C. (2022) ‘FOMC minutes sentiments and their impact on financial markets’, *Journal of Economics and Business*, 118, p. 106021. Available at: https://doi.org/10.1016/j.jeconbus.2021.106021.

Taylor, J.B. (1993) ‘Discretion versus policy rules in practice’, *Carnegie-Rochester Conference Series on Public Policy*, 39, pp. 195–214. Available at: https://doi.org/10.1016/0167-2231(93)90009-L.

Wooldridge, J.M. (2012) ‘Introductory Econometrics: A Modern Approach’.

Appendix

A. Text Pre-processing

1. Text Unclean

```
'Regarding the international outlook, signs of a faster reopening in China
and a less severe downturn in Europe eased concerns about global growth, co
ntributing to a depreciation in the exchange value of the dollar and suppor
ting optimism about emerging market economies.'
```

2. Text Clean

```
['regarding', 'international', 'outlook', 'signs', 'faster', 'reopening',
'china', 'less', 'severe', 'downturn', 'europe', 'eased', 'concerns',
'global', 'growth', 'contributing', 'depreciation', 'exchange', 'value',
'dollar', 'supporting', 'optimism', 'emerging', 'market', 'economies']
```

B. Ordered Probit Estimation

1. FinBERT Specification

finBERT Ordered Probit Results

	Dependent variable: $\Delta r_t + 1$		
	Naive Model	Sentiment	finBERT
	(1)	(2)	(3)
Δr_t	0.058***	0.045***	0.053***
	(0.009)	(0.009)	(0.009)
Sentiment$_t$		0.466***	
		(0.117)	
finBERT$_t$			0.319***
			(0.107)
-25bps	-1.566***	-1.729***	-1.628***
	(0.143)	(0.166)	(0.150)
+25bps	1.119***	1.215***	1.166***
	(0.066)	(0.073)	(0.069)
Observations	214	214	214
Pseudo R^2	0.18	0.24	0.21
AIC	225	209	217
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 6 – Ordered Probit Estimation with FinBERT Sentiment

2. Zero-Lower Bound Exclusion Specification

Ordered Probit Results (ZLB Exclusion)

	Dependent variable: $\Delta r_t + 1$		
	Naive Model	Sentiment	Hawkish/Dovish
	(1)	(2)	(3)
Δr_t	0.047^{***}	0.037^{***}	0.038^{***}
	(0.009)	(0.009)	(0.009)
$Sentiment_t$		0.503^{***}	
		(0.130)	
$Hawkish_t$			0.009^{***}
			(0.003)
$Dovish_t$			-0.009^{***}
			(0.002)
-25bps	-1.305^{***}	-1.358^{***}	-1.253^{***}
	(0.154)	(0.174)	(0.486)
+25bps	0.931^{***}	1.045^{***}	1.038^{***}
	(0.084)	(0.092)	(0.092)
Observations	141	141	141
Pseudo R^2	0.140	0.213	0.210
AIC	195.3	181.2	184.0
Note:		*p<0.1; **p<0.05; ***p<0.01	

Table 7 – Ordered Probit Estimation with ZLB Exclusion