Can We Extrapolate Future Yields From Central Banker Speeches?

Prottoya Chowdhury

May 8th, 2025

1 Introduction

Central bankers communicate their thoughts on the current and future trajectory of the economy through multiple different mediums. Being part of the central banking institution that overlooks various aspects of a nation's economy, these individuals are privy to non-public information that affects future central banking policy. Moreover, the personal opinions of these individuals and their tone towards various macroeconomic issues can reflect a consensus about future central banking policy that has not been explicitly communicated to the public.

In this paper, we seek to implement natural language processing (NLP) techniques to uncover latent information that is encoded within speeches delivered by central bankers both in an official setting, such as central banking policy initiatives, and unofficial settings, such as voluntary speaking events at colleges. Having uncovered this latent information, we then test whether it is predictive of yields across various term structures. Ultimately, on a global scale we saw Value of Currency associated with positive change in long term yields across a longer time horizon and domestic growth sentiment associated with a slight increase in shorter term yields after 3 months. For individual countries, we uncovered several interest insights. For the UK and Japan, it appears that domestic growth and currency value sentiment is predictive of changes in short term yields. In addition, for Australia, sentiment on trade balance appears to have a strong positive correlation with changes in short term yields over a year, with quite a large impact.

We believe the remarks of central bankers may contain two variants of latent information as follows:

1. **Personal Opinions on Current and Future Central Banking Policy**: Some central bankers possess more sway and influence than others. While addressing audiences, these individuals may integrate their opinions and thoughts on the current and future state of the economy. Note that these opinions voiced by these central bankers often find their way into central bank activity. Thus, uncovering indicators of these latent opinions that are present in the semantics of the speech texts is likely to convey information about future central banking activity. And that activity in turn should influence changes in yields across the term structure.

2. **Shared Consensus Amongst Central Banker**: Some central bankers obviously don't possess enouh influence to sway central banking activity significantly, but these individuals interact with a body of individuals at the central bank whose opinions together form the basis of current and future central banking activity. While information conveyed by these individuals is unlikely to directly influence yields or determine central banking activity, it can indicate a consensus amongst central bankers about future central banking activity. Thus, uncovering latent information encoded within speeches delivered by these individuals could be predictive of future central banking activity, which in turn would influence future yields across the term structure.

This study borrows inspiration from the realm of machine learning but grounds itself in macroeconomic theory. We divided our analysis into three different segments. We will describe each of them below.

The first segment entails dividing each speech in the corpus into relevant sections that correspond to a macroeconomic variable ('topic area') that affects lending rates. Recognizing that yields on debt can vary for various macroeconomic variables, such as trade balances, the value of the currency, inflation rates, domestic growth, we first sought to identify portions of each speech that aligned with each of the areas. One of our key assumptions was that sentiment across each of these areas might vary and that each of these macroeconomic issues might independently affect yields. Using a pre-trained Large Language Model to generate embeddings and then applying sentence similarity techniques, we divided each document on a sentence-level into these four different categories.

The second segments involves deriving sentiment scores and text-level embeddings for each of the macroeconomic issues for each of the speeches in the corpus. First, to derive sentiment scores, we used a pre-trained Large Language Model(LLM) that had been trained on a large dataset of financial texts. Using the outputs of the LLM, we then devised our own sentiment index, ranging from 0-30, with higher numbers indicating more positive sentiment. Second, to derive our text-level embeddings for each topic for each speech, we vectorized each division of our speech using the embedding model used in the first segment of the analysis.

Finally, the third part of our analysis involves regressing sentiment scores on yields across the term structure to assess the correlation of sentiment scores with short term and long term debt yields. By regressing topic level sentiment scores on yields, we hoped to uncover whether firstly, these sentiments were able to predict future yields across the term structure, and secondly, whether sentiments from certain topics had a stronger impact on yields than others. In addition, through regressing principal components of text-level embeddings for text corresponding to a topic area for each speech, we aimed to identify whether some latent vector space exists that could best describe the relationship between yields and the speeches delivered by central bankers.

One close parallel to this study is Petropoulos and Siakoulis (2021) [6] who use a corpus of 10,000 central bank speeches to predict episodes of financial turbulence, whereby the S&P500 drops by about 8% in a span of about 3 months. The authors segmented their analysis into two parts, whereby they first applied Extreme Gradient Boosting to develop a dictionary of words that labelled words in their corpus either positive of negative. Then using mutiple machine learning techniques, such as SVMs, Random Forests and Neural Networks, they created a sentiment index for each speech in their corpus.

However, the above study doesn't try to estimate correlation between central banker speeches and yields and nor does it ground itself in economic theory. As such, this study hopes to blend economic theory with machine learning techniques to produce more reliable and interpretable results.

Another study, Schmeling and Wager (2025) [7], uses a counting-based approach to estimate sentiment for each speech in their corpus. By counting the number of negative words based upon the Loughran and Mcdonald financial lexicon, they assume that the prevalence of more negative words should constitute a more negative sentiment for the document. While this is true to some extent, modern NLP techniques by using LLMs offer a significant advantage as they are able to generate sentiment scores by taking into account the entire context of a sentence or piece of text. Moreover, large LLMs have been trained on copious amounts of real world text, which means their training has already familiarized them with text similar to the ones in the speeches.
One study by Ahrens et al (2024) [1] closely resembles some of the strategies used in this study. Similar to our approach, Ahrens et al focus their analysis on speeches made by central bankers that are separate from official central bank communications. They develop a multi-modal NLP model that is trained on Federal Reserve Greenbook and strives to forecast changes in GDP and other macroeconomic variables that are one quarter ahead. They find that, contrary to popular belief where speeches by central bankers are believed to reduce market volatility, speeches by central bankers tend to increase tail-risk and asset price volatility. This lends some credence to our study's claim that central bankers, during the course of their speeches, reveal their own or shared consensus views on macroeconomic variables that can impact yields and hence prices on debt assets.

In addition, in a more recent paper by Hilscher et al (2024) [5], we see the application of NLP methods by using pre-trained neural networks. Hilscher et al make use of FinBert sentiment model, a neural network trained on large amounts of financial text, to assess the sentiment implied by official communications from the Federal Reserve and the European Central Bank. They find that the sentiment derived from the model was able to predict policy interest rates several quarters into the future and generally seemed to obey Taylor's Rule. Interestingly, they also find that equity market returns along with expectations for deflation tend to drive Federal Reserve sentiment, which would imply that the recent trajectory of public equity assets tends to drive how the Federal Reserve approaches setting and regulating interest rates in the economy. While our study does not account for returns and volatility within equity markets, this study elicits one of the core drivers for central banks to change interest rates. Thus, it shows us that central bankers seem to place important on recent equity returns and that these factors seem to influence their future actions which determine the trajectory of interest rates.

Examining central bank statements and central banker speeches and evaluating their impact on financial markets has long been a key area of research in quantitative finance. Bernanke and Kuttner (2004) [3] concluded that unexpected changes in monetary policy communicated through statements at monetary policy meetings could significantly influence expected excess returns on stocks, lending credibility to the idea that statements by central banker can both predict and impact financial markets. Moreover, from our prior findings of Bekaert et al (2013), it has been shown that monetary easing tends to decrease risk aversion and thereby lending rates [2].

While all these studies try to quantify impacts to financial markets as a result of central bank activity or commentary, many do not use available text data and base their theory sole on economic models. However, this study aims to improve upon these studies by using ideas in macroeconomics to inform its use of NLP techniques that can be used to uncover latent information present in non-standard mediums of data such as text. By using LLMs that have been trained on similar financial text and using various dimension reduction methods, the study tries to answer two questions, one from the perspective of mathematics ,and another from the perspective of economics:

1. Mathematial Question:What underlying vector space or dimensions best describe relationships between central bank speeches and yields across the term structure?

2. Economic Question:What particular macroeconomic variable and its sentiment is best able to predict yields across the term structure? And does this predictive power change as we try to predict yields across various maturities?

2 Data

To gather our speeches that are used to derive sentiment scores, we used a publicly avaialble dataset[4] on kaggle.com. The dataset includes 7721 speeches delivered by representatives from 8 central banks from the US, European Union, Sweden, Australia, Japan, Canada, Switzerland and UK. The dates of the speeches range from the late 1990s to about 2022. While most of the text has been cleaned in the dataset, there are some cases where the text for speeches was not cleaned properly (i.e. all periods had been removed which made it difficult to extract sentences from a speech). Our approach was to remove these speeches from our analysis, which happened to reduce our overall corpus of text to roughly 6000 speeches.

Second, to gather our dependent variables–yields–we used data from FRED. For each of the countries represented in our dataset of speeches, we collected historical yields on the 10-year government bond yield and the 3-month interbank lending rate. While we initially intended to collect data on various other yields across the term structure, it was difficult to find publicly available data on these yields that stretched back about 20 or so years. As such, we limited our analysis to purely these two yields. One improvement for this study would be to incorporate more intermediate term yields like yields on 1-year, 2-year and 5-year government and corporate bonds.

3 Generating Sentiment Scores

Instead of generating one generic sentiment for each document, we chose to generate sentiment scores for each speech across four macroeconomic variables : domestic growth, inflation, balance of trade and currency value.

To do so, we first split each document in our corpus into its constituent sentences. Formalizing this, we started with our vector of documents,$D = [X_1, X_2, \ldots, X_{6000}]$ and applied a function F such that :

$$= F(D) = [F(X_1), F(X_2), \ldots, F(X_{6000})]$$

$$= \begin{bmatrix} X_{1,1} & X_{1,2} & \ldots & X_{1,i} \\ \vdots & \vdots & \vdots & \vdots \\ X_{6000,1} & X_{6000,2} & \ldots & X_{6000,k} \end{bmatrix}$$

Where i, k denote the number of sentences within each document and $X_{a,b}$ represents the bth sentence in document a.

Next, we vectorized each of the sentences in the above representation by using the All Mini L6 V2 Embedding module, which is a neural network that takes in text with a maximum context length of about 500 tokens and produces a 384-dimensional vector representation for that group of tokens. Thus, we were able to produce a unique 384-dimensional vector for each sentence within a speech using this technique. To mathematically represent this, let L represent the function applied by our embedding module. For a sentence S within a document we have:

$$= L(S) = [s_1, \ldots, s_{384}]$$

Thus, for each document X_i for $i = 0, \ldots 6000$ within our corpus with k sentences in each document, we produced $k \times 384$ matrix.

The next goal we had was to classify each of the sentences into one of the four macroeconomic areas discussed above. Once, we were able to group all the sentence vectors for a document into the four different categories, we could use of the groups to generate a sentiment score for the category represented by the group. To go about classifying, we generated 100 key vectors using Open AI Chat GPT on each of the four different macroeconomic variables and vectorized them. Then, we took the mean for each component of the 384-dimensional vectors to arrive at a mean key vector that was representative of the topic area. Mathematically, this implies for a topic area A with sentences $A_1, \ldots, A_{100}$, we first produced the following embeddings:

$$= \begin{bmatrix} L(A_1) \\ L(A_2) \\ \vdots \\ L(A_{100}) \end{bmatrix}$$

$$= \begin{bmatrix} a_{1,1} & \cdots & a_{1,384} \\ \vdots & \vdots & \vdots \\ a_{100,1} & \cdots & a_{100,384} \end{bmatrix}$$

Then, we created a single 384-dimensional vector by taking the component-wise means of the above matrix. Thus we have:

$$= \text{Key Vector Per Macro Topic} = \begin{bmatrix} \frac{\sum_{i=1}^{100} a_{i,1}}{100} \\ \vdots \\ \frac{\sum_{i=1}^{100} a_{i,100}}{100} \end{bmatrix}'$$

Once, we had generated a mean key vector for each of the four topic areas, we computed the cosine similarity for each sentence vector per document with each of the four mean key vectors. Cosine similarity produces a value within $[-1,1]$ with a higher value associated with greater similarity between two vectors in an arbitrary vector space. The expression for computing cosine similarity is the following:

$$= \text{Cosine Similarity}(X,Y) = \frac{X \cdot Y}{\|X\| \times \|Y\|}$$

Thus for each sentence, we were able to generate 4 cosine similarity scores. We then classified the vector into the topic area with which it had the highest cosine similarity score. In this manner, we were able to group all the sentence vectors in each document into the four different macroeconomic variables.

Finally, to generate the sentiment scores per category, we took each group of vectors per document per topic area and computed the average vector. Thus for k vectors in a document for topic a topic area j, we computed the component-wise mean to arrive at a mean vector for that topic per document.

Next, we treated these mean topic area vectors as input into the Finlang sentiment score model, which produced a softmax probability output across three different sentiments : positive, neutral and negative. To compute our sentiment scores per topic area, we used the following function:

$$= \text{Topic Sentiment Score} = \begin{cases} 10 \times P(negative) & if \;\; P(negative) \geq P(positive) + P(neutral) \\ 20 \times P(neutral) & if \;\; P(neutral) \geq P(positive) + P(negative) \\ 30 \times P(positive) & if \;\; P(positive) \geq P(neutral) + P(negative) \end{cases}$$

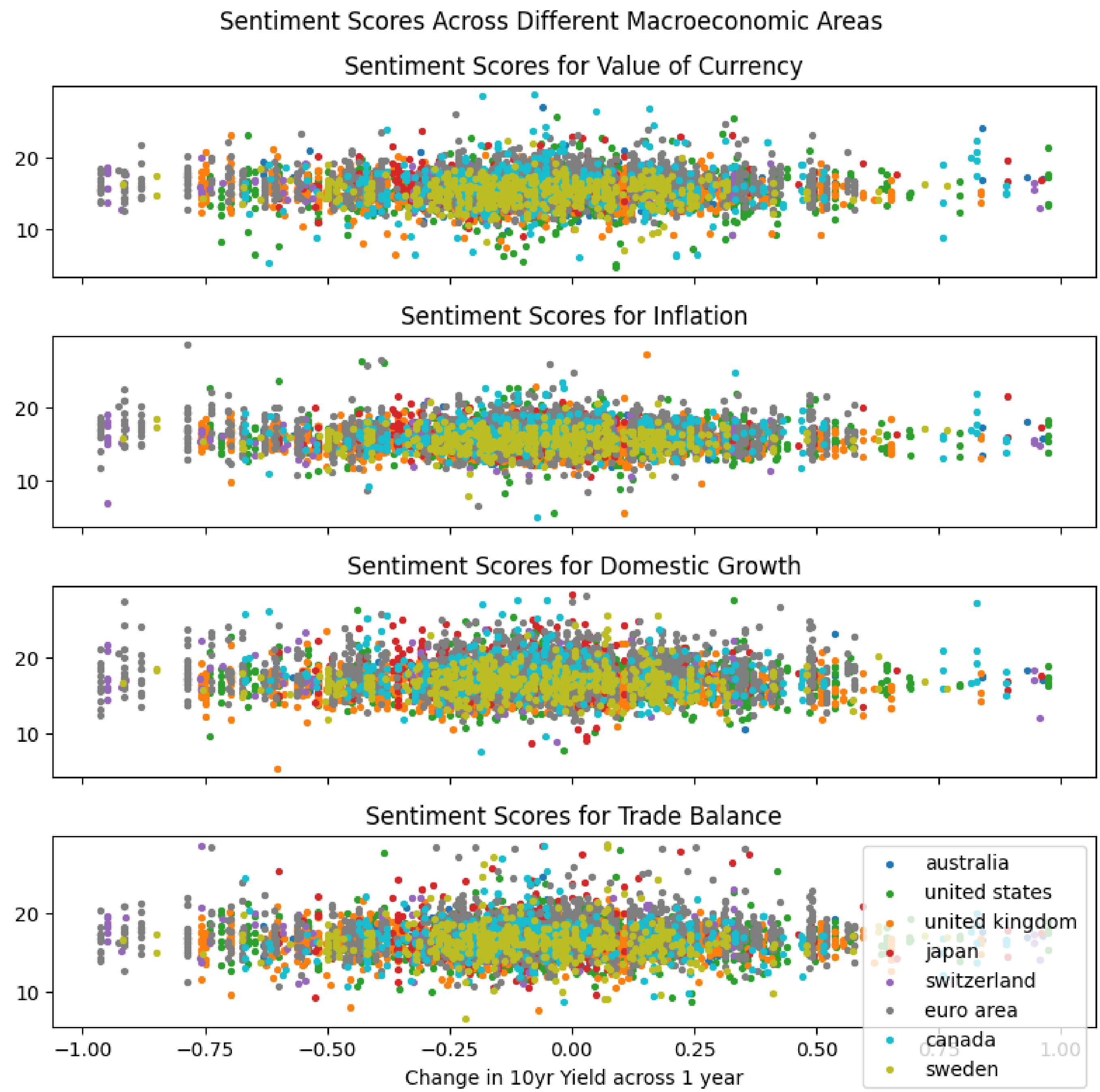
For each topic area per document, this function produced a setniment score between [0,30]. The descriptive statistics for our sentiment scores are shown below: Also, below,

Descriptive Statistics for Sentiment Scores

	Inflation	Domestic Growth	Trade Balance	Value of Currency
count	5797.000000	5797.000000	5797.000000	5797.000000
mean	15.968372	17.259118	16.769701	15.868556
std	1.773048	2.270728	2.094848	2.102164
min	4.876917	5.417383	6.682375	4.772609
25%	14.871137	15.805927	15.621077	14.649099
50%	15.922433	17.092604	16.697930	15.926356
75%	16.931645	18.484864	17.803231	17.088057
max	28.762434	28.419263	28.626338	28.660192

Table 1: The following table provides summary statistics for the sentiment scores. Note that while we initiall started with an initial corpus of 7721 speeches, we had to remove some of the speeches from our analysis as they weren't cleaned properly. Hence, we eliminated all speeches that we were unable to extract individual sentences from for our analysis

we have provided a scatterplot illustration for our sentiment scores across all countries and speeches in our data set.



4 Model Setup

Since our study aimed to find the association between changes in debt yields and speech sentiment, our model was trying to predict the changes in debt yields for various time lags with sentiment scores. We used 1-month, 2-month, 3-month, 6-month and 12-month changes in yields of the 10-year government bonds and 3-month interbank lending rates as our dependent/prediction variable for our model. Letting $Y_{k,t}$ be prediction variable where $k = 1, 2, 3, 6, 12$ and denotes the change in yields for a debt security and t denotes the time that we start measuring the change in yields, we had the following model :

$$= Y_{k,t} = \alpha_{k,t} + \beta_{1,k} \times \gamma_{inflation,t} + \beta_{2,k} \times \gamma_{currency,t} + \beta_{3,k} \times \gamma_{trade,t} + \beta_{4,k} \times \gamma_{growth,t} + \epsilon_{k,t}$$

$$\text{where} \quad Y_{k,t} = \frac{\text{Yield at time t+k} - \text{Yield at time t}}{\text{Yield at time t}}$$

In the model above, $\gamma_{inflation,t}$ is our sentiment score for inflation on a speech at time t, $\gamma_{currency,t}$ is our computed sentiment score for currency value at time t,$\gamma_{trade,t}$ is our computed sentiment score for balance of trade at time t and $\gamma_{growth,t}$ is our sentiment score for domestic growth at time t.

The unknown variables $\beta_{i,k}$ for $i = 1, 2, 3, 4$ denotes the percentage change in yields of a debt instrument that is associated with an increase of the sentiment score by 1.

5 Results

We now discuss the results our study using the model described in the previous section. First, we conducted the regression as specified in the previous section on the entire set of our speeches to see if any of the macroeconomic variables was predictive of yield changes across our entire sample. The results are shown below.

Significant Regression Results for Sentiments Across All Countries

Table 2:

	twelve_month_10y	three_month_3m_ib
const	-0.0792959*	-0.1188863***
	(0.0434053)	(0.0309258)
Inflation	-0.0031773	-0.0006341
	(0.0021990)	(0.0015667)
Domestic Growth	-0.0019400	0.0030601**
	(0.0017711)	(0.0012619)
Trade Balance	0.0009164	0.0000049
	(0.0018779)	(0.0013380)
Value of Currency	0.0044085**	0.0032943**
	(0.0018437)	(0.0013136)
R-squared	0.0013696	0.0029934
R-squared Adj.	0.0006800	0.0023048
N	5797	5797

Standard errors in parentheses.
* p¡.1, ** p¡.05, ***p¡.01

Table 3: This table only shows the regressions where we calculated some of the unknown parameters to be significant. In total, we performed about 9 different regressions using our model across two yields and 5 different time lags. It appears that sentiment on economic variables is predictive of shorter term rates on shorter time horizons and longer term debt on longer time horizons. That said, note that the regression paramaters are quite small so the effect is quite muted.

When we look at the entire dataset, it appears that a single unit increase in the sentiment for value of currency correlated with 0.4% increase in yield of 10-year government bonds. For shorter term debt, in this case the 3-month interbank rate, a 1 point increase in the sentiment score for Domestic Growth correlates with 0.3% increase in yields and a 1-point increase in the sentiment for value of currency correlates with a 0.3% increase in yields. Note that the effect sizes are quite small in this case. This is to be expected as, across the different countries represented in our dataset, some economies are more or kess affected by some of the economic variables. As such, while we do get some small significant results through our model, the correlation is certainly not very strong.

Because comparing changes in yields for all the countries in our dataset produces some rather noisy results, we decided to run our regression model on groups of speech sentiment scores for each country. These results show that different macroeconomic variable sentiments have different levels of correlation with interest rates across various countrys. We discuss some of the more interesting results below.

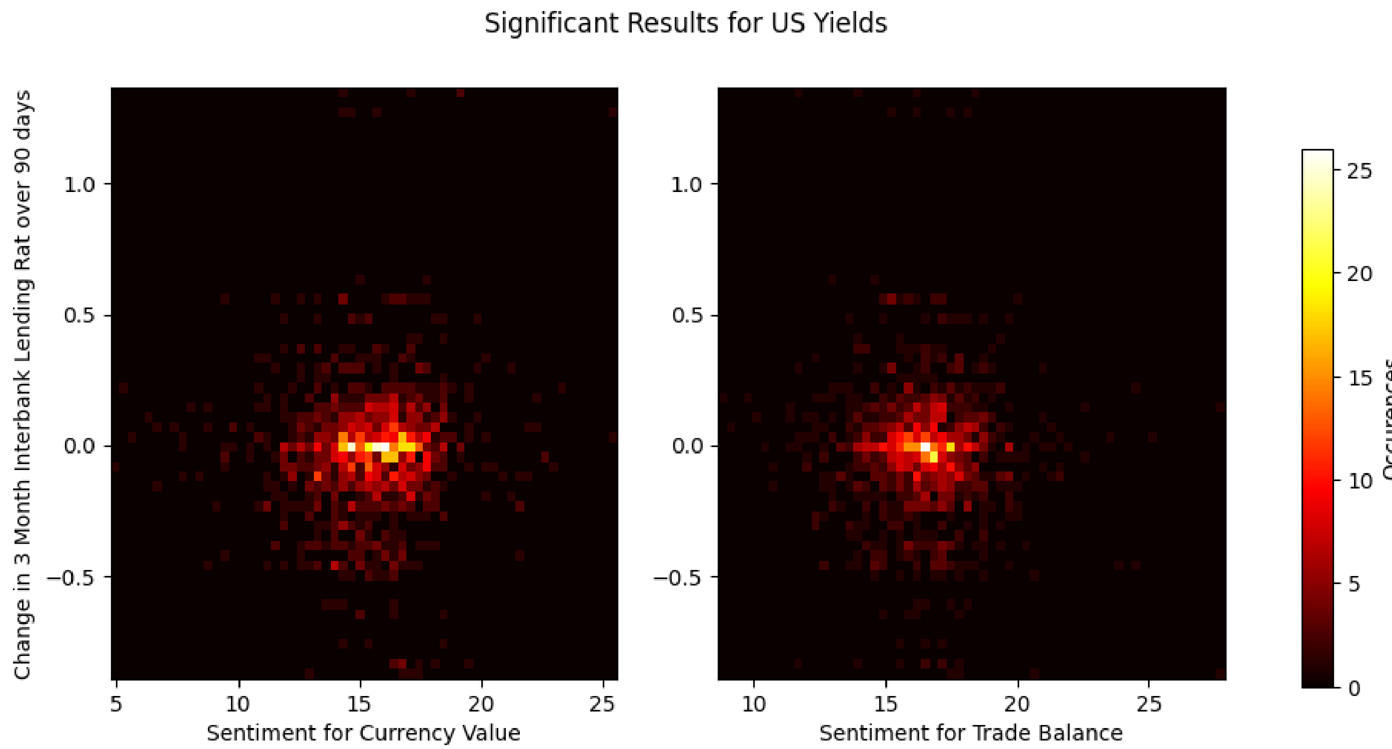
First, we present the results of our regression model on only U.S central bank speeches within our dataset along with a graphical illustration.

Significant Results for US Central Bank Sentiments

Table 4:

	$3_{mi}b$
Inflation	-0.0033036
	(0.0037422)
Domestic Growth	0.0009442
	(0.0034386)
Trade Balance	-0.0065453*
	(0.0034801)
Value of Currency	0.0076027**
	(0.0030968)
R-squared	0.0190562
R-squared Adj.	0.0157672
N	1197

Standard errors in parentheses.
* p¡.1, ** p¡.05, ***p¡.01



In the case of the US, sentiment on Value of Currency seems to have a significant relationship with the 3-month interbank lending rate. A 1 point increase in that sentiment correlates with about a 0.8% increase in yields for the 3-month interbank rate. However, if we look at the heatmap for changes in yields across sentiment scores, we see that the relationship is still quite noisy. There seems to be a lot of sentiment scores that are near 15, which would indicate a neutral sentiment. This could indicate either that our NLP methods to generate sentiment scores were insufficient at identifying the latent information within speeches from US central bankers, or it could signal that US central bankers typically exercise greater caution when delivering their speeches to ensure their tone is neutral.

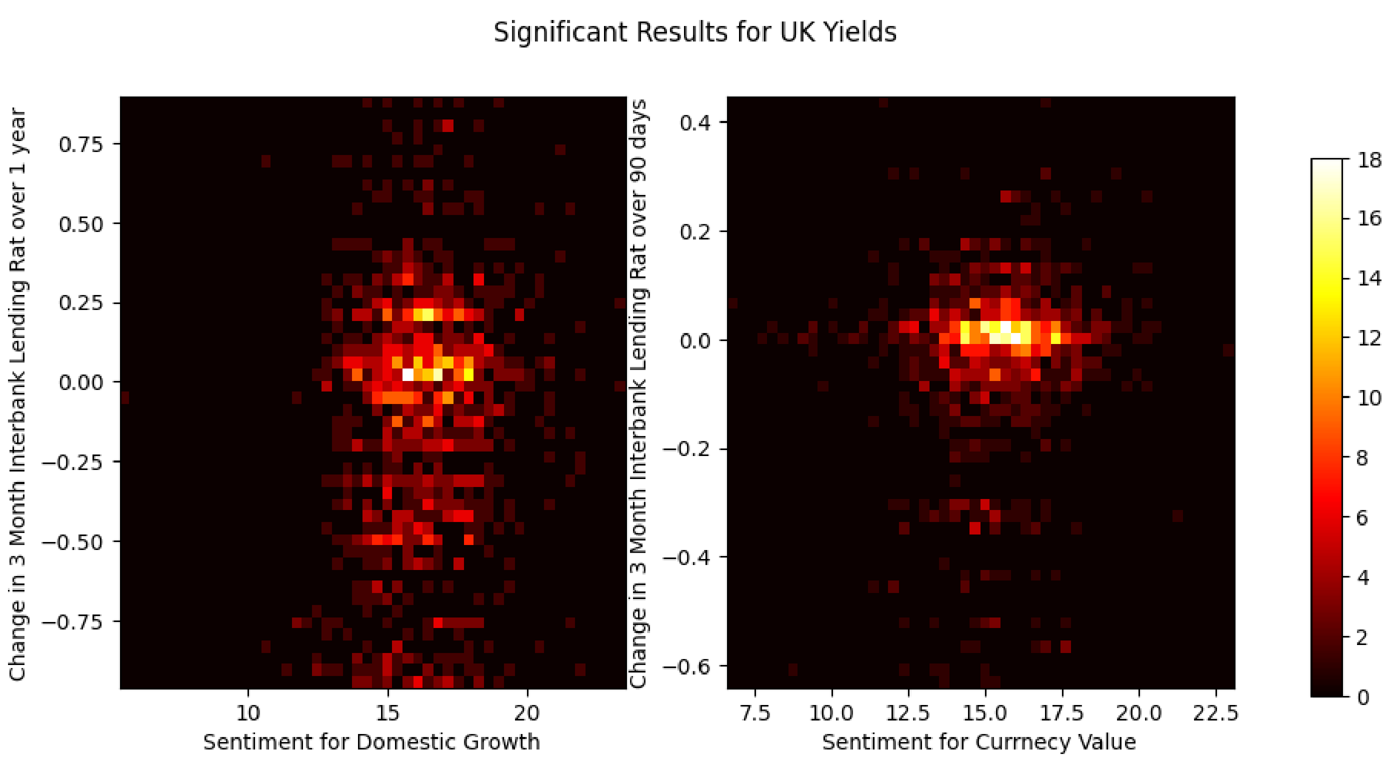
Now, we will examine our results for UK central bank speeches and sentiment scores.

Significant Results for UK Central Bank Sentiments

Table 5:

	three_month_3m_ib	twelve_month_3m_ib
const	-0.2154094***	-0.3029285*
	(0.0616503)	(0.1600515)
Inflation	-0.0011989	-0.0124140
	(0.0031554)	(0.0081917)
Domestic Growth	0.0053196*	0.0210618***
	(0.0027185)	(0.0070575)
Trade Balance	0.0024091	0.0087733
	(0.0029157)	(0.0075695)
Value of Currency	0.0052402**	-0.0056160
	(0.0026549)	(0.0068923)
R-squared	0.0130660	0.0123427
R-squared Adj.	0.0092032	0.0084771
N	1027	1027

Standard errors in parentheses.
* p¡.1, ** p¡.05, ***p¡.01



In the case of the UK, a 1 point increase in the sentiment on Value of Currency is associated with 0.5% increase in 3-month interbank yields over the next 3-months. However, we do find a more significant result if we extend our time lag to 12 months. Over 1 year, it seems that 1 point increase in the sentiment for Domestic Growth is associated with 2.1% increase in yields. This shows that the actions of central banks in the UK might be motivated to a significant degree by concerns over domestic growth.

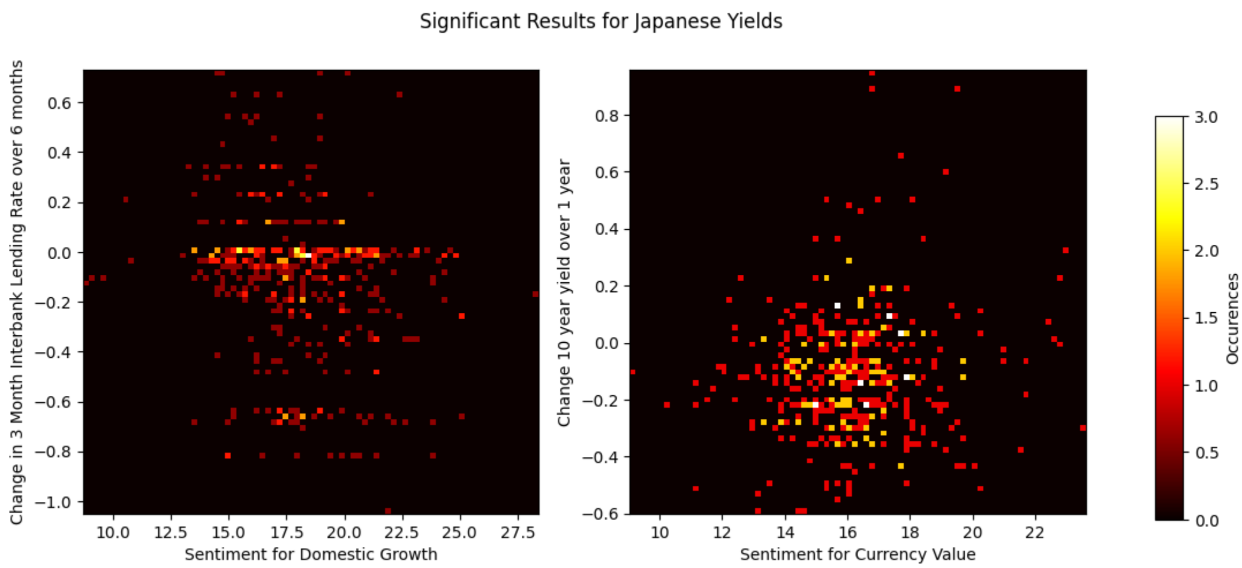
Next, we present our results for Japan.

Significant Results for Japanese Central Bank Sentiments

Table 6:

	twelve_month_10y	six_month_3m_ib
const	-0.3170747***	0.1433929
	(0.1161748)	(0.1587041)
Inflation	0.0065926	0.0008019
	(0.0068082)	(0.0093005)
Domestic Growth	-0.0057537	-0.0229586***
	(0.0046305)	(0.0063256)
Trade Balance	0.0027518	-0.0033759
	(0.0044663)	(0.0061013)
Value of Currency	0.0096077*	0.0129599*
	(0.0057144)	(0.0078063)
R-squared	0.0135012	0.0460174
R-squared Adj.	0.0038768	0.0367103
N	415	415

Standard errors in parentheses.
* p¡.1, ** p¡.05, ***p¡.01



In the case of Japan, we see that the sentiment on value of currency and Domestic Growth seem to be associated strongly with changes in future 3-month inter-bank lending rates. Specifically, a 1 point increase in the sentiment score for Domestic Growth is associated with a -2.3% change in future 6 month yields for 3-month inter-bank rates. This does seem to agree with generall macroeconomic theory as central banks that want to encourage domestic growth seek to increase investment by lowering policy rates. This then allows businesses and households to take on greater amounts of debt, leading to more investment spending and consumption in the economy. In addition, the Japanese central bank has in the past expressed interest in maintaining a stable and strong yen, and we some of that reflected in the results above.

In the plots, we clearly a negative correlation between 3-month interbank lending rates and sentiment scores for Domestic growth and a somewhat positive relationship between currency-value and the change in the yield of 10 year Yen bond across one year. This further validates our results for Japan.

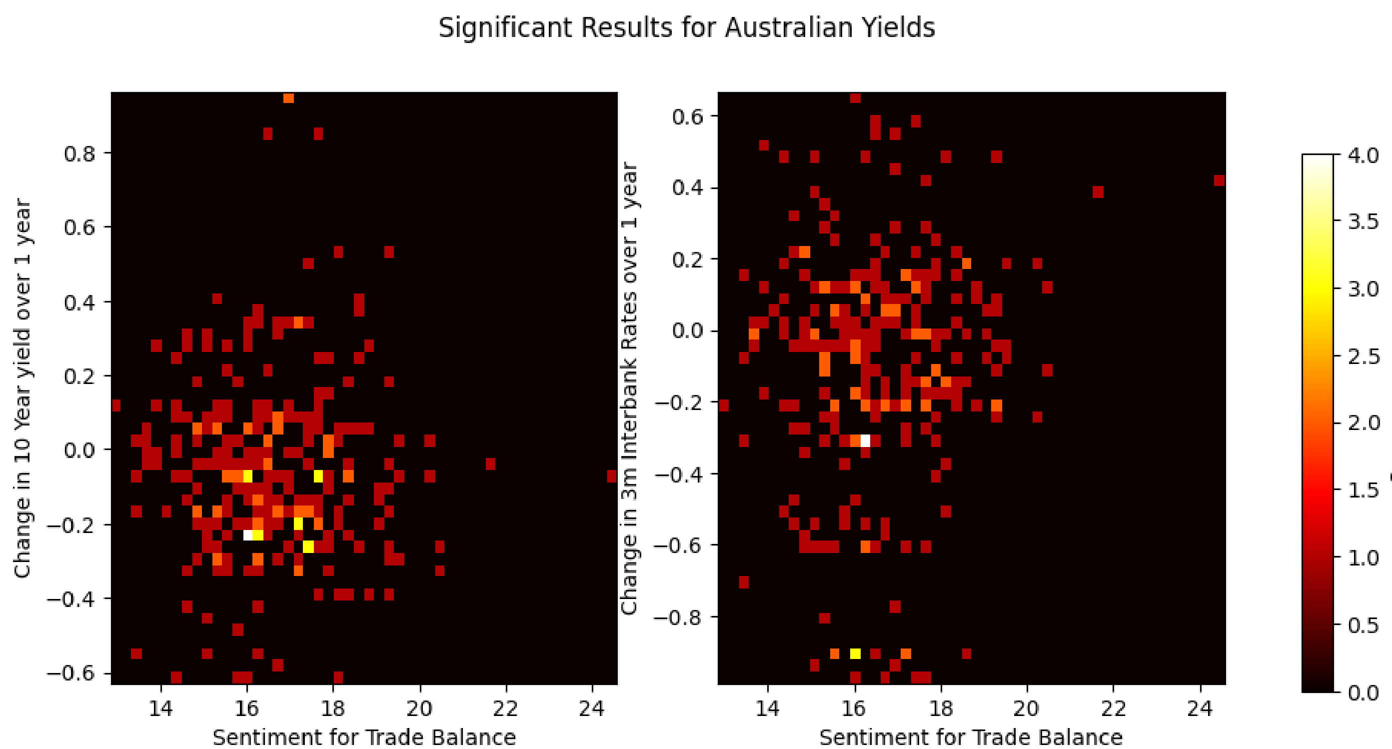
Next, we will examine our results for Australia.

Significant Results for Australian Central Bank Sentiments

Table 7:

	twelve_month_10y	twelve_month_3m_ib
const	0.2236570	0.0076318
	(0.2463239)	(0.3300548)
Inflation	-0.0141258	-0.0241290
	(0.0116351)	(0.0155901)
Domestic Growth	-0.0119478	-0.0070441
	(0.0097241)	(0.0130295)
Trade Balance	0.0042990	0.0298450**
	(0.0102347)	(0.0137137)
Value of Currency	0.0048848	-0.0075695
	(0.0078201)	(0.0104783)
R-squared	0.0145752	0.0263738
R-squared Adj.	-0.0010047	0.0109805
N	258	258

Standard errors in parentheses.
* p¡.1, ** p¡.05, ***p¡.01



In the case of Australia, sentiment on trade balance appeared to have a strong positive correlation with changes in the 3-month interbank rate across 1 year. A 1 point increase in the Sentiment score for trade balance is associated with a 3% increase in the 3-month interbank rate across 12 months. This aligns with Australia's economy which is heavily reliant on the exports of coal, metals and other natural resources.

6 Conclusion

This study aimed to find latent information embedded within speeches delivered by central bankers across different countries and see if that information correlates with yields across the term structure. On a global level, we saw that the sentiment surrounding the Value of Currency was highly correlated with the 1-year change yields for 10-year government debt. In addition, we saw that sentiment on domestic growth was correlated with a slight increase in yields for 3-month interbank rates after 3 months. However, both of these results, while significant, were very small. This was mainly due to the different structures of economies that were part of our dataset.

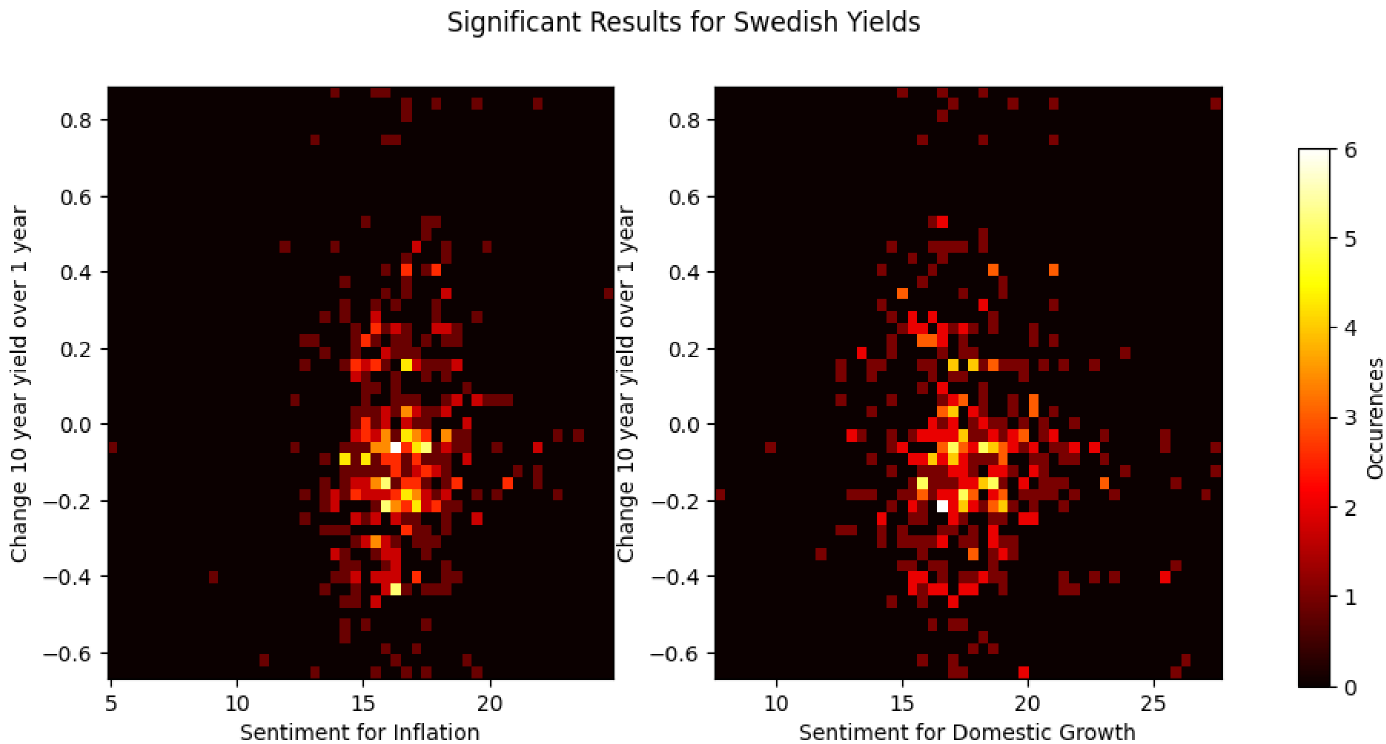
A more honest and insightful approach that we took involved examining the relationship of sentiment scores across the macroeconomic topics and yields for each specific country. We found some significant results, especially in the case of Australia and Japan. However, we did suffer from smaller sample sizes for both Australia and Japan and would need more data to further validate our results.

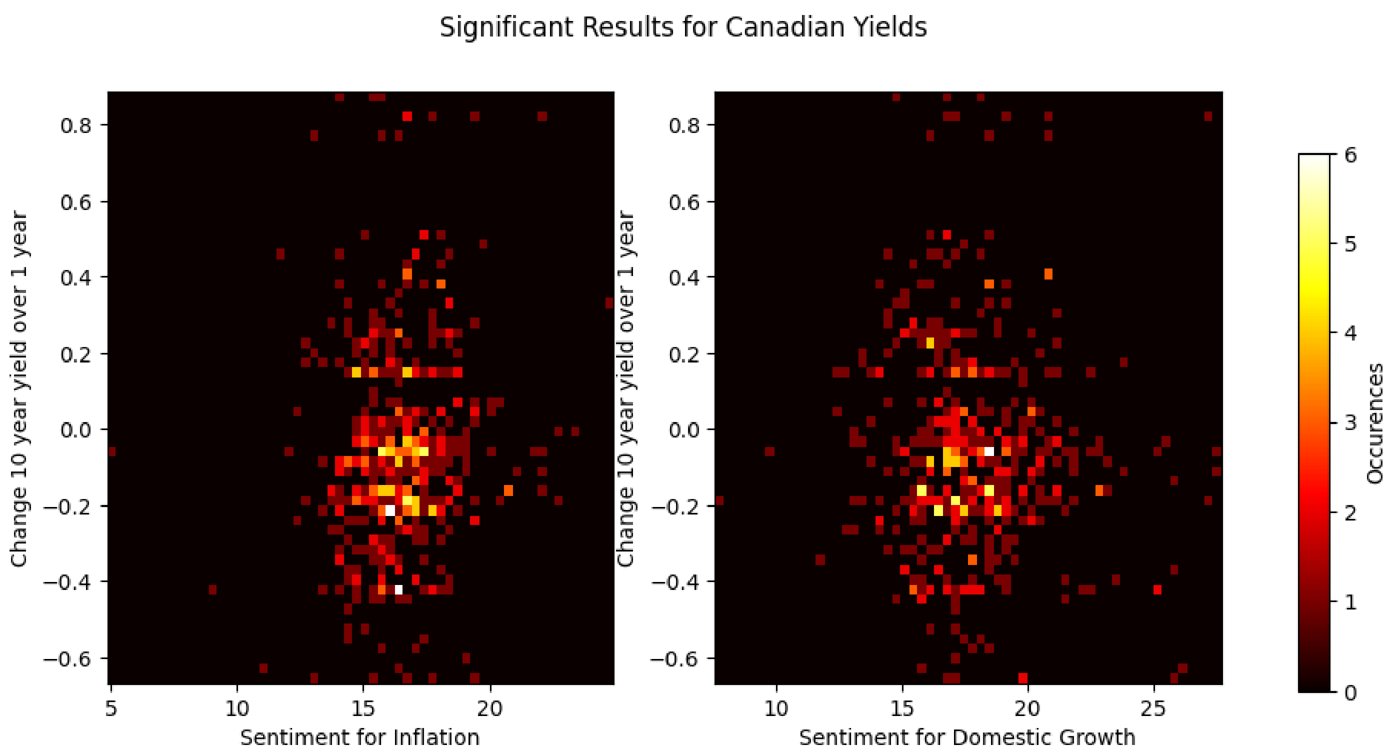
On the whole, this study sought to blend modern NLP techniques with ideas from macroeconomic theory to carry out an analysis using novel information hidden within text data. Our core premise was that there exists some latent information that reflects either the direct opinions or consensus shared among central bankers representing a central banking institution. Given that we showed some of our results as being statistically significant, this assumption appears to be true. It is either the case that the latent information expressed by central bankers through their speeches directly affects central bank activity and policy rates or that central bankers,through the speeches they deliver, highlight a shared consensus for future action by the central bank. Our study is not able to confirm which of these two channels might be more or less correct and we would need more granular data to be able to identify through which channel these speeches influence central bank policies.

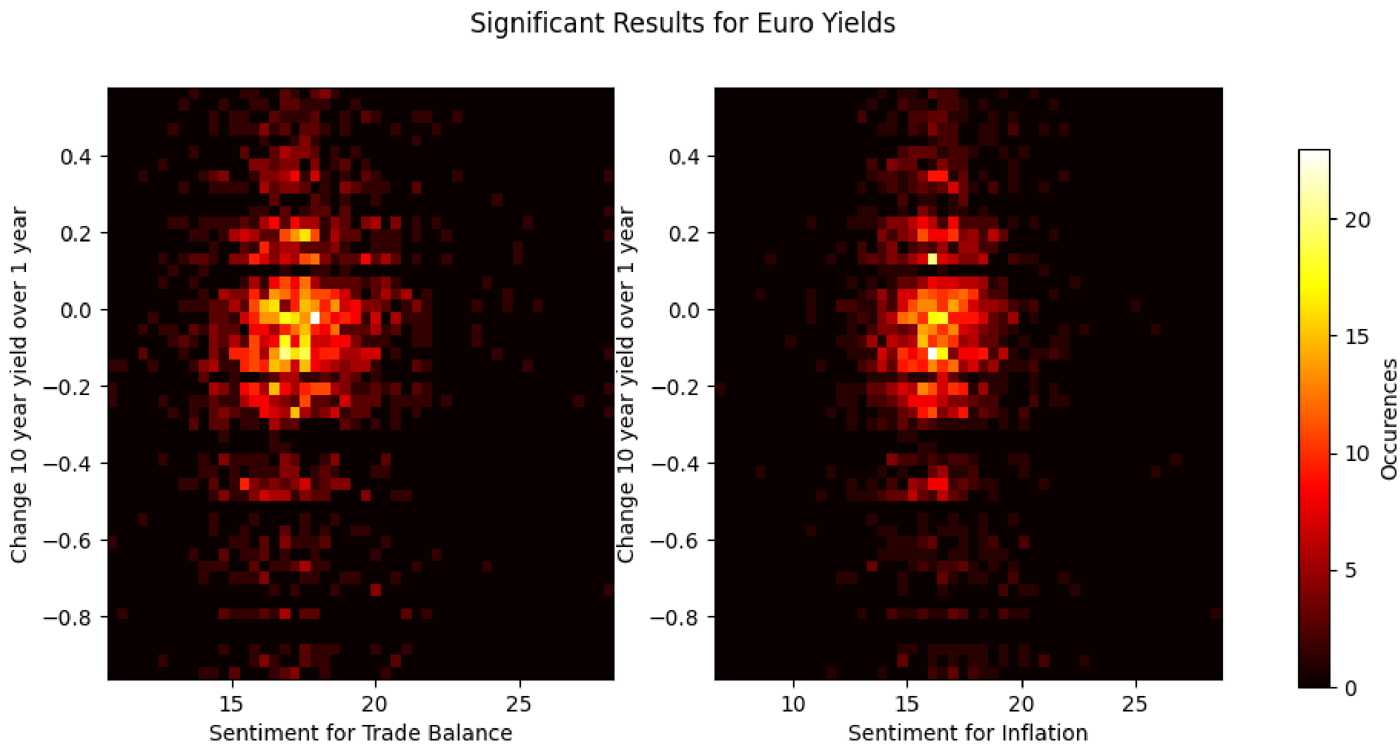
One of the clear drawbacks of our approach is the inability to derive causality. It seems unlikely that a single speech by one central banker would cause the central bank to undertake actions that change yields across the term structure. In order to derive causality, we propose examining central bank activity following the speeches delivered by prominent central bankers whose opinions are most likely to shape monetary policy within a given economy.

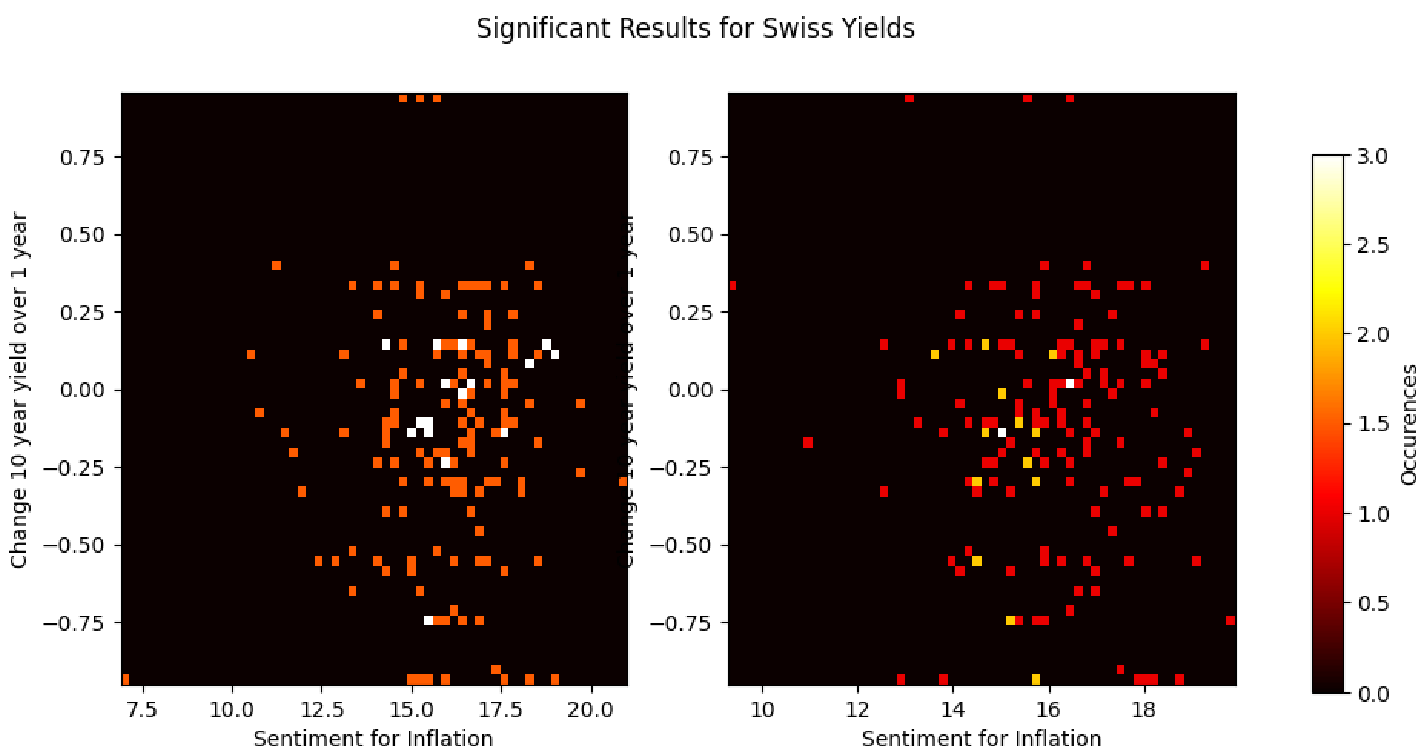
Another drawback of our study stems from the lack of compute power. We used rather simple embedding models to vectorize our speeches and derive sentiment scores, and it is likely that these simpler models were unable to fully extract or identify the latent information within central banker speeches. Furthermore, some of the pre-trained models we used such as the All Mini l6 V2 embedding model were trained on general text and not financial text. We recommend that for future studies, a researcher either tries to fine tune these models across a larger corpus of central bank speeches and communications, or that they use an embedding model that has been specifically trained on central bank speeches and communications. This would enable the models to identify more latent information and could thus lead to greater clarity for results.

7 Appendix and Models Used









Models Used:

1. FinLang/financial-embeddings-investopedia
2. all-MiniLM-L6-v2

References

[1] Maximilian Ahrens, Deniz Erdemlioglu, Michael McMahon, Christopher J Neely, and Xiye Yang. Mind your language: Market responses to central bank speeches. *Journal of Econometrics*, page 105921, 2024.

[2] Geert Bekaert, Marie Hoerova, and Marco Lo Duca. Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60(7):771–788, 2013.

[3] Ben S Bernanke and Kenneth N Kuttner. What explains the stock market's reaction to federal reserve policy? *The Journal of finance*, 60(3):1221–1257, 2005.

[4] David Gauthier. Central bank speeches. 2022.

[5] Jens Hilscher, Kyle Nabors, and Alon Raviv. Information in central bank sentiment: An analysis of fed and ecb communication. *Available at SSRN 4797935*, 2024.

[6] Anastasios Petropoulos and Vasilis Siakoulis. Can central bank speeches predict financial market turbulence? evidence from an adaptive nlp sentiment index analysis using xgboost machine learning technique. *Central Bank Review*, 21(4):141–153, 2021.

[7] Maik Schmeling and Christian Wagner. Does central bank tone move asset prices? *Journal of Financial and Quantitative Analysis*, 60(1):36–67, 2025.