Does FED Communication cause immediate and abnormal returns stock market?

A sentiment analysis on the S&P 500 Stock Market Index

Matthias Farngruber Matriculation Number 1 Filip Lukijanovic 11776896
Peter Prlleshi 11776041

Contents

Abstract

Literature Review

The following research project is mostly based on the work done by Möller and Reichmann (2021) in the field of sentiment analysis. In their paper "ECB Language and Stock Returns - A Textual analysis of ECB Press Conferences" they explore the impact of the language used by the ECB in their regular press conferences on stock returns in the Euro Area.

They achieve this by mining the statements and ranking the sentiments of each statement based on general tone, uncertainty and constraint. Once this is classified, the researchers then cross-check the high frequency intraday data for Euro Area stock returns on each statement day, by employing the technique of an event study. This allows them to see how stock returns reacted to statements by the ECB at whatever point they wanted to examine them.

As mentioned, the researchers focused on tone, uncertainty and constraining language for the sentiment analysis. Tone captures the overall language - or the overarching narrative - of a statement, uncertainty measures how ambiguous a statement may be and constraining language quantifies how constraining the ECB communicates to be in the future. Of course the researchers did not read through every single statement by the ECB, instead, they employed a dictionary-based sentiment mining approach that considered grammatical and syntactical cues to analyze the sentiment expressed in ECB press conferences. Afterwards they scored each statement with regards to each category by employing heuristic-adjusted sentiment scores based on word lists used in previous studies. Möller and Reichmann (2021a) The Authors had used a plethora of controls in their regression analyses, an approach that we were more reluctant to follow due to our slightly smaller sample size. The main reason behind that is that our sample could hardly accommodate the sheer number of variables used in the intitial analysis. Below, we delve into the subset of data used, and into our methodology for modelling Möller and Reichmann's (2021) sentiment methodology. In addition, we explore whether transformer-based approaches that are bespoke to Federal Reserve Statements can outperform heuristics-adjusted sentiment approaches using dictionaries, i.e. our proprietary approach. We do so by investigating the explanatory power of regressions using each sentiment approach.

Data

The data used for this project consists of the following:

- Federal Reserve (FED) Meetings: The U.S. American FED holds regular meetings multiple times per year where they talk about the current economic situation and what the plan is going forwards. This leaves us with 74 statements over 10 years from January 30th 2013 until July 26th 2023.
- Standard & Poor's 500 stock market index pricing: We cross-check the statements and their nature with the price of the S&P 500 stock market index on each day of a statement. This index is of particular interest since it incorporates 500 U.S. american companies representing a large part of the whole market itself. This means that we can use the S&P 500 as a proxy for how the market behaves at a given day. Alongside that we obtained weighted debt to equity ratios of the index to see if higher leverage leads to stronger reactions to sentiment. A third and final element of S&P 500 data was lagged returns that were used in some of our analysis
- Federal Funds Futures data: the rate or price equivalent of the interest rate that the market prices in for the next 30 days was obtained as well, as a proxy for market expectations on interest rates.

All of our data, save for the FOMC meeting minutes were obtained via the Bloomberg Terminal data services, whereas the FOMC minutes were manually extracted.

Methodology

Our analysis of equity returns following FOMC sentiment levels can be partitioned in two. The first part is sentiment extraction, wit our approach closely following that of Möller and Reichmann (2021), as well as an implementation of Google's FinBERT NLP transformer for an alternative gaugue of the tone of FOMC minutes. The second part focuses on the construction of the variables needed in our regression, both core variables as are equity returns and sentiment, as well as various controls inspired by the paper from Möller and Reichmann.

Sentiment Extraction - VADER Method

Our first approach to modelling the sentiment of FOMC minutes closely resembles that of Möller and Reichmann (2021), who use the NLTK VADER package in Python to extract sentiment. The VADER package is special insofar as it adjusts sentiment for heuristics, which gives it an advantage over conventional dictionary-based NLP methods (Hutto, 2014) Hutto and Gilbert (2014). Furthermore, sentiment can be modeled for intensity as well, with Sentiment $S \in [-4, 4]$. Möller and Reichmann (2021) expand the inbuilt dictionary of words and corresponding stand-alone sentiment scores by financial terms that correspond to mere Tone, Uncertainty, as well as Constraining language, so as to be able to capture different dimensions of speech. MISSINGMISSING: \

Library Construction for each sentiment CON UNC, i.e. References to the libraries via BIBTEX«

Our approach to this was to update the lexicon in VADER as well, whereby our extraction of Tone merely necessitated the updating of the pre-existing lexicon. For Unc and Con, however, we used the libraries by Bodnaruk (2015) Bodnaruk, Loughran, and McDonald (2015) for constraining language and Loughran and McDonald (2011) Loughran and McDonald (2011) to construct scores for uncertain language and for extending the VADER dictionary by financial terms. To extract the nuances of uncertain and constraining language, the sentiment scores of all tokens not included in the respective external libraries received reduced sentiment scores, to 20% of their initial sentiment score. Meanwhile, the tokens from the libraries were weighted regularly, i.e. with sentiment scores $S \in \{2, -2\}$, as 2 is equivalent to moderate intensity in the VADER package. Adjusting every word's sentiment score manually for its intensity was not deemed feasible for the scope of a project.

Once these transformations took place, the FOMC data was tokenized into sentences, which were then tokenized for each constituent word. The sentiment score for every sentence is weighted by the word count in each sentence relative to the total word count in the statement.

```
simple_tone <- lm(abnormal_return ~ Tone, data = Core)
simple_unc <- lm(abnormal_return ~ Unc, data = Core)
simple_con <- lm(abnormal_return ~ Con, data = Core)
simple_bert <- lm(abnormal_return ~ Bert, data = Core)

stargazer(simple_tone, simple_con, simple_unc, simple_bert, column.labels = c("Tone", "Unc", "Con", "Be
```

Table 1:

	Dependent variable:			
	abnormal_return			
	Tone	Unc	Con	Bert
	(1)	(2)	(3)	(4)
Tone	0.013			
	(0.010)			
Con		0.009		
		(0.024)		
Unc			0.018	
			(0.019)	
Bert				0.001
				(0.043)
Constant	−0.002	−0.001	−0.003	−0.00003
	(0.002)	(0.003)	(0.004)	(0.041)
Observations	76	76	76	76
R^2	0.023	0.002	0.011	0.00000
Adjusted R^2	0.010	−0.011	−0.002	−0.014
Residual Std. Error (df = 74)	0.010	0.010	0.010	0.010
F Statistic (df = 1; 74)	1.774	0.156	0.859	0.0002

Note: *p<0.1; **p<0.05; ***p<0.01

```r
simple_tone_IR <- lm(abnormal_return~Tone + IR, data = Core)
simple_unc_IR <- lm(abnormal_return~Unc + IR, data = Core)
simple_con_IR <- lm(abnormal_return~Con + IR, data = Core)
simple_bert_IR <- lm(abnormal_return~Bert + IR, data = Core)

stargazer(simple_bert_IR, simple_bert_IR, simple_bert_IR, simple_bert_IR,
          column.labels = c("Tone", "Unc", "Con", "Bert"), header = F)
```

Table 2:

	Dependent variable:			
	abnormal_return			
	Tone	Unc	Con	Bert
	(1)	(2)	(3)	(4)
Bert	0.009	0.009	0.009	0.009
	(0.045)	(0.045)	(0.045)	(0.045)
IR	−0.001	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	−0.007	−0.007	−0.007	−0.007
	(0.043)	(0.043)	(0.043)	(0.043)
Observations	76	76	76	76
R^2	0.006	0.006	0.006	0.006
Adjusted R^2	−0.021	−0.021	−0.021	−0.021
Residual Std. Error (df = 73)	0.010	0.010	0.010	0.010
F Statistic (df = 2; 73)	0.211	0.211	0.211	0.211
Note:	*p<0.1; **p<0.05; ***p<0.01			

```r
simple_tone_LR <- lm(abnormal_return ~ Tone + IR + lagged_return, data = Core)
simple_unc_LR <- lm(abnormal_return ~ Unc + IR + lagged_return, data = Core)
simple_con_LR <- lm(abnormal_return ~ Con + IR + lagged_return, data = Core)
simple_bert_LR <- lm(abnormal_return ~ Bert + IR + lagged_return, data = Core)

stargazer(simple_tone_LR, simple_con_LR, simple_unc_LR, simple_bert_LR, column.labels = c("Tone", "Unc"
```

```r
simple_tone_SUR <- lm(abnormal_return ~ Tone + IR + lagged_return + Surprise, data = Core)
simple_unc_SUR <- lm(abnormal_return ~ Unc + IR + lagged_return + Surprise, data = Core)
simple_con_SUR <- lm(abnormal_return ~ Con + IR + lagged_return + Surprise, data = Core)
simple_bert_SUR <- lm(abnormal_return ~ Bert + IR + lagged_return + Surprise, data = Core)

stargazer(simple_tone_SUR, simple_unc_SUR, simple_con_SUR, simple_bert_SUR, column.labels = c("Tone", "
```

```r
simple_tone_DE <- lm(abnormal_return ~ Tone + IR + lagged_return + Surprise + debt_equity, data = Core)
simple_unc_DE <- lm(abnormal_return ~ Con + IR + lagged_return + Surprise + debt_equity, data = Core)
simple_con_DE <- lm(abnormal_return ~ Con + IR + lagged_return + Surprise + debt_equity, data = Core)
simple_bert_DE <- lm(abnormal_return ~ Bert + IR + lagged_return + Surprise + debt_equity, data = Core)

stargazer(simple_tone_DE, simple_unc_DE, simple_con_DE, simple_bert_DE, column.labels = c("Tone", "Unc"
```

Table 3:

	Dependent variable:			
	abnormal_return			
	Tone	Unc	Con	Bert
	(1)	(2)	(3)	(4)
Tone	0.018			
	(0.014)			
Con		−0.001		
		(0.034)		
Unc			0.016	
			(0.020)	
Bert				0.010
				(0.046)
IR	0.001	−0.001	−0.0003	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)
lagged_return	−0.090	−0.060	−0.067	−0.062
	(0.162)	(0.163)	(0.162)	(0.162)
Constant	−0.003	0.001	−0.002	−0.008
	(0.004)	(0.005)	(0.004)	(0.043)
Observations	76	76	76	76
R^2	0.030	0.007	0.016	0.008
Adjusted R^2	−0.011	−0.034	−0.025	−0.034
Residual Std. Error (df = 72)	0.010	0.011	0.010	0.011
F Statistic (df = 3; 72)	0.736	0.172	0.382	0.188
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 4:

	Dependent variable:			
	abnormal_return			
	Tone	Unc	Con	Bert
	(1)	(2)	(3)	(4)
Tone	0.024*			
	(0.014)			
Unc		0.012		
		(0.020)		
Con			0.013	
			(0.034)	
Bert				0.015
				(0.045)
IR	0.0004	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)
lagged_return	−0.047	−0.023	−0.018	−0.019
	(0.160)	(0.162)	(0.162)	(0.162)
Surprise	0.022**	0.017	0.019*	0.018*
	(0.011)	(0.011)	(0.011)	(0.011)
Constant	−0.005	−0.001	−0.001	−0.013
	(0.004)	(0.004)	(0.005)	(0.043)
Observations	76	76	76	76
R^2	0.084	0.049	0.046	0.046
Adjusted R^2	0.033	−0.004	−0.007	−0.008
Residual Std. Error (df = 71)	0.010	0.010	0.010	0.010
F Statistic (df = 4; 71)	1.635	0.919	0.862	0.855
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 5:

	Dependent variable:			
	abnormal_return			
	Tone	Unc	Con	Bert
	(1)	(2)	(3)	(4)
Tone	0.032**			
	(0.015)			
Con		0.025	0.025	
		(0.037)	(0.037)	
Bert				0.003
				(0.050)
IR	−0.0003	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)
lagged_return	−0.057	−0.019	−0.019	−0.016
	(0.159)	(0.162)	(0.162)	(0.163)
Surprise	0.022**	0.019*	0.019*	0.017
	(0.011)	(0.011)	(0.011)	(0.011)
debt_equity	0.00003	0.00002	0.00002	0.00001
	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Constant	−0.016**	−0.008	−0.008	−0.005
	(0.008)	(0.009)	(0.009)	(0.045)
Observations	76	76	76	76
R^2	0.113	0.057	0.057	0.051
Adjusted R^2	0.050	−0.010	−0.010	−0.017
Residual Std. Error (df = 70)	0.010	0.010	0.010	0.010
F Statistic (df = 5; 70)	1.792	0.853	0.853	0.756

Note: *p<0.1; **p<0.05; ***p<0.01

```r
interactions_tone <- lm(abnormal_return ~ Tone * debt_equity + IR + lagged_return + Surprise * debt_equ

interactions_unc<-lm(abnormal_return~Unc*debt_equity + IR + lagged_return + Surprise*debt_equity + debt_

interactions_con<-lm(abnormal_return~Con*debt_equity + IR + lagged_return + Surprise*debt_equity + debt_

interactions_bert <- lm(abnormal_return~Bert*debt_equity + IR + lagged_return + Surprise*debt_equity + d

stargazer(interactions_tone, interactions_unc, interactions_con, interactions_bert, column.labels = c("
```

```r
trial_tone<-lm(abnormal_return~Tone + IR + lagged_return + Surprise + debt_equity, data=Core)
trial_unc<-lm(abnormal_return~Unc + IR + lagged_return + Surprise + debt_equity, data=Core)
trial_con<-lm(abnormal_return~Con + IR + lagged_return + Surprise + debt_equity, data=Core)
trial_bert<-lm(abnormal_return~Bert + IR + lagged_return + Surprise + debt_equity, data=Core)

stargazer(trial_tone, trial_unc, trial_con, trial_bert,
          column.labels = c("Tone", "Unc", "Con", "Bert"), header = F)
```

BMA

After regressing everything and getting the results we do a final check via Baysian Model Averaging, where we can see which of the Variables from our Core data set truly are important to abnormal returns. This methodology is to check which variables have variable importance in a regression.

```r
# Defining the Model
abnormal_return_model <- Core$abnormal_return ~ Core$Tone + Core$Unc + Core$Con + Core$Bert + Core$IR +

# Doing the BMA itself
bms_results <- bms(abnormal_return_model)
```

```
##                          PIP     Post Mean      Post SD Cond.Pos.Sign Idx
## Core$Surprise      0.05896313  9.967645e-04 4.782526e-03    1.00000000   7
## Core$Tone          0.05623271  1.074809e-03 5.605947e-03    1.00000000   1
## Core$Unc           0.02703853  4.553171e-04 4.441457e-03    0.96803485   2
## Core$debt_equity   0.02421109  2.725401e-07 3.423660e-06    0.95062231   8
## Core$Con           0.02401279 -1.773257e-04 7.184673e-03    0.69543177   3
## Core$IR            0.02300608 -1.155113e-05 1.646098e-04    0.08399735   5
## Core$lagged_return 0.02023056 -1.327445e-03 2.449888e-02    0.00000000   6
## Core$Bert          0.01926924  1.154585e-04 6.683659e-03    0.88214338   4
##
## Mean no. regressors               Draws             Burnins                Time
##            "0.2530"               "256"                 "0"   "0.06427097 secs"
##  No. models visited      Modelspace 2^K           % visited         % Topmodels
##               "256"               "256"               "100"               "100"
##            Corr PMP            No. Obs.         Model Prior             g-Prior
##                "NA"                "76"        "random / 4"               "UIP"
##     Shrinkage-Stats
##          "Av=0.987"
##
## Time difference of 0.06427097 secs
```

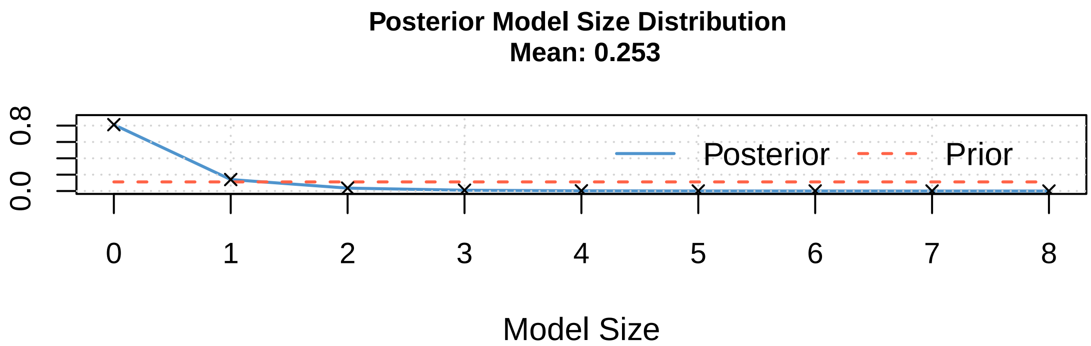
Table 6:

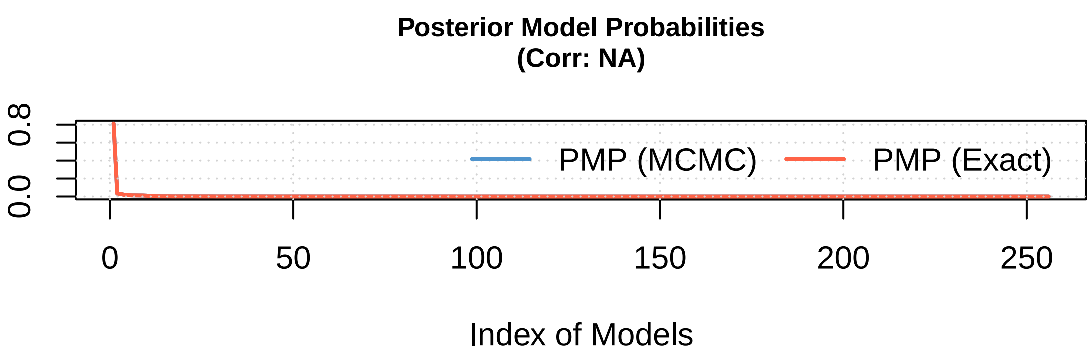
	Dependent variable:			
	abnormal_return			
	Tone	Unc	Con	Bert
	(1)	(2)	(3)	(4)
Tone	0.082			
	(0.053)			
Unc		0.143		
		(0.094)		
Con			0.147	
			(0.126)	
Bert				−0.095
				(0.207)
debt_equity	0.0001	0.0001	0.0001	−0.0002
	(0.00004)	(0.0001)	(0.0001)	(0.001)
IR	−0.001	−0.001	−0.001	−0.001
	(0.002)	(0.001)	(0.002)	(0.001)
lagged_return	−0.154	−0.166	−0.116	−0.086
	(0.158)	(0.163)	(0.161)	(0.165)
Surprise	0.305***	0.356***	0.310***	0.279***
	(0.102)	(0.107)	(0.103)	(0.103)
Tone:debt_equity	−0.0001			
	(0.0002)			
Unc:debt_equity		−0.0003		
		(0.0003)		
Con:debt_equity			−0.0003	
			(0.0004)	
Bert:debt_equity				0.0002
				(0.001)
debt_equity:Surprise	−0.001***	−0.001***	−0.001***	−0.001**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Constant	−0.032**	−0.041*	−0.029	0.082
	(0.015)	(0.022)	(0.019)	(0.193)
Observations	76	76	76	76
R^2	0.205	0.190	0.158	0.135
Adjusted R^2	0.123	0.106	0.071	0.046
Residual Std. Error (df = 68)	0.010	0.010	0.010	0.010
F Statistic (df = 7; 68)	2.500**	2.272**	1.819*	1.514
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 7:

	Dependent variable:			
	abnormal_return			
	Tone	Unc	Con	Bert
	(1)	(2)	(3)	(4)
Tone	0.032**			
	(0.015)			
Unc		0.024		
		(0.023)		
Con			0.025	
			(0.037)	
Bert				0.003
				(0.050)
IR	−0.0003	−0.002	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)
lagged_return	−0.057	−0.029	−0.019	−0.016
	(0.159)	(0.162)	(0.162)	(0.163)
Surprise	0.022**	0.015	0.019*	0.017
	(0.011)	(0.011)	(0.011)	(0.011)
debt_equity	0.00003	0.00002	0.00002	0.00001
	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Constant	−0.016**	−0.011	−0.008	−0.005
	(0.008)	(0.009)	(0.009)	(0.045)
Observations	76	76	76	76
R^2	0.113	0.066	0.057	0.051
Adjusted R^2	0.050	−0.001	−0.010	−0.017
Residual Std. Error (df = 70)	0.010	0.010	0.010	0.010
F Statistic (df = 5; 70)	1.792	0.990	0.853	0.756

Note: *p<0.1; **p<0.05; ***p<0.01





Bibliography

Bodnaruk, Andriy, Tim Loughran, and Bill McDonald. 2015. "Using 10-k Text to Gauge Financial Constraints." *Journal of Financial and Quantitative Analysis* 50 (4): 623–46.

Hutto, Clayton, and Eric Gilbert. 2014. "Vader: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text." In *Proceedings of the International AAAI Conference on Web and Social Media*, 8:216–25. 1.

Loughran, Tim, and Bill McDonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *The Journal of Finance* 66 (1): 35–65.

Möller, Rouven, and Doron Reichmann. 2021b. "ECB Language and Stock Returns – a Textual Analysis of ECB Press Conferences." *The Quarterly Review of Economics and Finance* 80 (May): 590–604. https://doi.org/10.1016/j.qref.2021.04.003.

———. 2021a. "ECB Language and Stock Returns – a Textual Analysis of ECB Press Conferences." *The Quarterly Review of Economics and Finance* 80 (May): 590–604. https://doi.org/10.1016/j.qref.2021.04.003.