Department of Economics, University of Warwick

Monash Business School, Monash University

as part of

Monash Warwick Alliance

Analysing the response of U.S. financial market to the Federal Open Market Committee statements and minutes based on computational linguistic approaches

Xuefan Pan

Warwick-Monash Economics Student Papers

March 2023 No: 2023-43

ISSN 2754-3129 (Online)

The Warwick Monash Economics Student Papers (WM-ESP) gather the best Undergraduate and Masters dissertations by Economics students from the University of Warwick and Monash University. This bi-annual paper series showcases research undertaken by our students on a varied range of topics. Papers range in length from 5,000 to 8,000 words depending on whether the student is an undergraduate or postgraduate, and the university they attend. The papers included in the series are carefully selected based on their quality and originality. WM-ESP aims to disseminate research in Economics as well as acknowledge the students for their exemplary work, contributing to the research environment in both departments.

"*We are very happy to introduce the Warwick Monash Economics Student Papers (WM-ESP). The Department of Economics of the University of Warwick and the Economics Department at Monash University are very proud of their long history of collaboration with international partner universities, and the Monash Warwick Alliance reflects the belief in both Universities that the future will rely on strong links between peer Universities, reflected in faculty, student, and research linkages. This paper series reflects the first step in allowing our Undergraduate, Honours, and Masters students to learn from and interact with peers within the Alliance.*"

Ben Lockwood (Head of the Department of Economics, University of Warwick) and Michael Ward (Head of the Department of Economics, Monash University)

Recommended citation: Pan, X. (2023) Analysing the Response of U.S. Financial Market to the Federal Open Market Committee Statements and Minutes based on Computational Linguistic Approaches. *Warwick Monash Economics Student Papers* 2023/43.

[1] Warwick Economics would like to thank Lory Barile, Gianna Boero, and Caroline Elliott for their contributions towards the selection process.

Analysing the response of U.S. financial market to the Federal Open Market Committee statements and minutes based on computational linguistic approaches

Pan Xuefan*

Abstract

I conduct content analysis and extent the existing models of analysing the reaction of the stock market and foreign currency markets to the release of Federal Open Market Committee (FOMC) statements and meeting minutes. The tone changes and uncertainty level of the monetary policy communication are constructed using the dictionary-based word-count approach at the whole document level. I further apply the Latent Dirichlet Allocation (LDA) algorithm to investigate the different impacts of topics in the meeting minutes. High-frequency data is used as the analysis is an event study. I find that the tone change and uncertainty level have limited explanation power on the magnitude of the effect of the release of FOMC documents especially statements on the financial market. The communication from FOMC is more informative for the market during the zero lower bound period, compared to the whole sample period.

Keywords: Monetary policy,Communication,Text Mining
JEL Classification: E52,E58

*Contact Email:pxfrico@gmail.com.
Online Appendix: https://github.com/SeptSwansea/Dictionary-based-word-count-content-analysis-.
Acknowledgement:I would like to thank Dr Fatih Kansoy for offering advice and sharing data with me. I also would like to thank Giri Rabindranath for sharing the coding tutorial of topic modelling on highdemandskills website. I also receive help from The Notre Dame Software Repository for Accounting and Finance which shares textual analysis resources. Many thanks to those warm-hearted respondents on the Stack Overflow website. Without them, this work can not possibly be done.

1 Introduction

Communication has gradually become a powerful tool for the central bank as expectations management helps improve the efficiency of the monetary policy (Blinder et al. 2008). Communication becomes particularly important when the short-term interest rate approaches 0 because there is no room for the interest rate to go down any further. The central bank stimulates the aggregate demand by promising to maintain the low policy rate for quarters. The expectation of the short-term interest rate will affect the long-term interest rate and consequently affects consumption and investment activities (Dincer and Eichengreen 2013).

After highlighting the motivation for increasing transparency of the monetary policy, I introduce the main ways of FED communication. FOMC statement is the announcement that attracts the most attention. The meeting minute is also in the toolkit and its impact increased after the adjustment of its release time at the end of 2004 (Danker and Luecke 2005). The effect of the release of the FOMC statements and minutes on the financial market is widely studied. To analyse the effect of statements on stock markets and foreign currency markets, I first apply Gürkaynak, Sack, and Swanson (2004)'s (GSS) model which decomposes the monetary policy shock into target surprise and path surprise using the updated data. The downside of using proxies is that they do not directly measure the information received by the market Hansen, McMahon, and Prat (2018). The wording of the central bank communication is considered to be a factor that affects the influence of the monetary policy (Gürkaynak, Sack, and Swanson 2004). Following this clue, I apply the word-count approach which allows us to quantify the information about feature policies received by the market which is supposed to be captured by the path surprise.

The effect of FOMC minutes relies on relative informativeness compared to the FOMC statements. I first examine the effect of the net tone and uncertainty level of the whole document, which measure the information conveyed through text materials. Macroeconomic indicators-interest rate, unemployment rate and a dummy variable which indicates the recession periods are added in the regression as well. The average length of minutes is longer than the one of statements, which allows us to apply LDA algorism and investigate the effect at the topic level (Jegadeesh and Wu 2017). In the second stage, the result of the LDA model is used to explore the effect of different topics which is a more segmented analysis compared to the first one.

In my results, there is no strong relationship between the sentiment scores of FOMC communication and the reaction of the market. However, during the zero lower bound period, the FOMC communication becomes more important and there are more statistically significant parameters in the regression compared to the one in the whole sample period. Thirdly, the stock market is paying more attention to the FOMC communication compared to the foreign currency markets. The fourth finding is that some topics in the FOMC meeting minutes are more informative to the market than others.

This paper is organized as follows. The second section introduces relevant lit-

erature. The third section is about methodology. Sections 3 and 4 are about data source and data description, respectively. Section 5 shows the empirical analysis result and discussion of the findings. In section 6, I list the shortcomings of this paper and where I can make improvements. Section 7 is all about the conclusion.

2 Literature Review

2.1 The introduction of monetary policy and policy communication

One of the main duties of the Federal Reserve is to implement the monetary policy to stabilize the price level and reduce the unemployment rate. Communication is one important tool. Central bank communication is defined as the release of information from the central bank to the public. There are four main aspects of the information the central bank is trying to convey. Policy objectives are the first part. The second part is about what actions the central bank has taken. Apart from discussing what has been done, the central bank also provides information on economic outlooks and offer a forecast for future policy rate (Blinder et al. 2008).

Regarding how the central bank communicates to the market, the fed releases policy documents after the announcement of the policy actions. Meeting statements which are published shortly after the meetings are the one that attracts the most public attention (Tadle 2022). The second important way of communication is through meeting minutes. The distinct features of these two meeting documents lie in the amount of information contained and the releasing timings (Rosa 2013). The meeting minutes are published three weeks after the meeting and offer a detailed description of what has been discussed in the meeting, views from different committee members, policy actions and the reason behind and voting outcomes (Danker and Luecke 2005). In this paper, I only include pre-scheduled communication from the committee and exclude communication from committee members like Fed chairman speeches.

2.2 Why communication matters

The mechanism by which the communication from the central bank affects the economy is that it can influence market expectations of interest rates (Blinder et al. 2008).

Blinder et al. (2008) classify studies on central bank communication into two categories. The studies diverge on the channels of how central bank communication affects expectations. One strand of the researchers thinks that the central bank has more insight into the economic outlook and that communication brings new information to the market. C. D. Romer and D. H. Romer (2000) examine whether the federal reserve does have superior information to the public and find the federal reserve has a better prediction of inflation as well as other real output indicators which indicates that information asymmetry exists. Nakamura and Steinsson (2018) refer to this "Fed information effect" and introduce a model

to explain the effect of monetary policy. In this new model, the information from the central bank not only has an impact on expected inflation but also can affect the market's anticipation of fundamental economic indicators which further affects economic activities. For example, a tightening policy can create an optimistic atmosphere which leads to more investment and consumption.

The other category of the study focuses on how the increasing transparency of the central bank action reduces the uncertainty in the market which is related to the risk-based channel. Schmeling and Wagner (2019) find that the release of statements of the European central bank can affect the price of volatility insurance which is related to the risk premia of the asset. The change of yield difference among corporate bonds with different risk levels is also found to be related to the tone of the statements.

Communication from the central bank becomes particularly important during the zero lower bond period. Unconventional monetary policies-forward guidance and quantitative easing are designed to cope with the situation where the federal fund rate is approaching 0 and the central bank can't stimulate the economy by further lowering the interest rate (Rudebusch et al. 2018).

2.3 Communication measurement

The influence of both statements and minutes is explored in this paper. Many papers have discussed the effect of statements on asset prices. Gürkaynak, Sack, and Swanson (2004) introduce the main structure of analysing the impact of releasing statements on asset prices. They show the change in the federal fund rate alone can not fully explain the whole variation of the asset prices after the announcement. The "future path of policy" is added as the second surprise factor. The second surprise here is to capture information about future policy paths. Hausman and Wongswan (2011) further introduce two proxies for these two factors to examine the influence of the statements on asset prices. The downside of using proxies, as Hansen, McMahon, and Prat (2018) point out, is that it does not directly reflect what information is conveyed. Many papers have emphasized that the wording of the statements matters. Nakamura and Steinsson (2018) suggest the content of the information conveyed can determine the causal effect of monetary surprises. Gürkaynak, Sack, and Swanson (2004) also states the role of the phrasing of FOMC statements in affecting the asset price when the target federal fund rate has been predictable these years. The computational linguistic approach comes in to quantify the content of monetary policy communication and helps fill this gap.

The basic structure of analysing the effect of statements on the financial market in the GSS paper mentioned above can not apply to the analysis of minutes because a part of the information especially policy actions is revealed to the public through statements three weeks ago and it is digested by the market already (Nechio, Wilson, et al. 2016). Nechio and Wilson test whether the tone change of minutes relative to statements affects treasury bond yield and emphasize that it is the relative change that brings the impact. Tadle (2022) shares a similar idea and mentions that if the minute doesn't bring new information about the economic outlook, investors will not change their expectations formed

three weeks ago. Relative informativeness is the key to the specification of the model. The informativeness can be measured by the computational linguistic approach.

In summary, the content analysis by using computational linguistic approaches offers a new perspective in analysing the effect of central bank communication. Researchers have applied different methods to quantify the information. Jansen and De Haan (2007) create an indicator which is about the use of the word 'vigilance' in the European Central Bank communication to measure the ECB's expectation of risk. Although they find limited explanation power of this indicator. Apel and Grimaldi (2014) apply the search-and-count-words approach to measure the tone of the meeting minutes and find that the tone change in the minutes can help predict the monetary policy of the Swedish Central Bank. Boukus and Rosenberg (2006) use the Latent Semantic Analysis model to further explore the effect of different topics in FOMC meeting minutes on the market. Latent Dirichlet Allocation is used by Jegadeesh and Wu (2017) to examine whether the tones of FOMC meeting minutes are correlated to the market reaction.

2.4 Market reaction measurement

The reaction of the stock market to monetary policy is studied by many researchers. The reason why many researchers focus on the reaction of asset prices is that they can subsequently affect real economic activities. The response of the financial market to the monetary policy is difficult to measure. Monetary policies are designed to cope with shocks, which means the effect of policy actions and financial shocks are interacting with each other, making it hard to isolate one another (Nakamura and Steinsson 2018). Many papers use the Vector Autoregression(VAR) model to cope with the confounding variables problem (Rudebusch et al. 2018; Hansen and McMahon 2016). High-frequency, event-study analysis is another method to mitigate the problem which is used by Rosa (2013), Hausman and Wongswan (2011), Cieslak and Schrimpf (2019) and many others.

The global impact of the Fed monetary policy is another field of interest because the USA is a large open economy. Rosa (2011) finds that the exchange rates react prominently to the policy actions and communication from the Fed. Pinchetti and Szczepaniak (2021) also find that there is an international spillover of the Fed announcement. Investors will adjust their portfolios and capital flows into markets with lower risk when there is a contradictory monetary policy.

The main aim of this paper is to decompose the effect of monetary policies on financial markets. My models are supplements to the existing models of analysing FOMC statements and minutes by implementing content analysis with the help of a computational linguistic approach. It helps us understand the transmission mechanism of monetary policy. The equity market and the foreign currency market are the two financial markets of interest.

3 Methodology

3.1 The reason for conducting evet-study analysis and using high-frequency data

Fawley, Neely, et al. (2014) point out there are omitted variables bias and simultaneity bias in the analysis of asset prices response to monetary policies which will lead to inconsistent estimation results. To better understand the two biases, I quote the linear system they introduced in their paper.

$$\triangle f_t = a_0 + a_1 \triangle R_t + a_2 news_t + \varepsilon_{f,t}$$

$$\triangle R_t = b_0 + b_1 \triangle f_t + b_2 news_t + \varepsilon_{R,t}$$

f_t represents policy change and R_t is asset price change. They affect each other and move simultaneously. $news_t$ is other macro information that can affect monetary policy and asset prices. We can see from the equation system that the absence of other related macro news variables and the simultaneous movement between monetary policy and asset price can cause biases when we apply OLS. The solution is to narrow down the event window which leaves no time for policymakers to adjust the policy and reduces the possibility of other news being released at the same observation time interval.

I conduct event-study analysis, which is one standard way of measuring asset price response to some announcement (Binder 1998). The 30-minute window is set to be around the announcement time (13:50-14:20) following the setting of Hausman and Wongswan (2011) which effectively avoid the effect of other macro information on the asset price. Jegadeesh and Wu (2017) confirmed most other important economic news is not released during this time on Wednesday and Thursday. The high-frequency data is used to capture the change in asset prices in a small range of time intervals.

3.2 The basic analysis for FOMC statements

The specification of the regression is based on Hausman and Wongswan (2011)'s work.

$$R_t = \alpha + \beta_1 TS_t + \beta_2 PS_t + \varepsilon_t$$

As we can see from the equation, the variation of the asset price is composed of two parts. The first part-target surprise captures the unanticipated Fed fund rate changes, which equals the revealed target policy rate minus the anticipated fed fund rate. The second element is path surprise, which is unexpected changes in the projected future rate.

Proxies are designed to represent the expected level of the policy rate. Kuttner (2001) first uses a contract-based Fed funds future price in a market at the Chicago Board of Trade as a proxy for the expected federal fund rate. This future rate reflects the average effective Fed funds rate in a calendar month. Hamilton (2009) further examines the statistical property of daily fluctuations of the future price and concludes it is a good indicator for market expectations of the monetary policy. There are certain features of futures where the advantages lie in. As summarized by Kansoy (2019), CME offers a futures market where

information is easily transmitted with low transaction costs. Besides, daily market-to-market prices make the daily update of price changes in the futures possible. CME also limits the credit risk for market participants. In addition, liquid tools are available for investors to adjust investment strategies in response to monetary policy change.

After the discussion of the property of the proxy for market expectations, I quote an example to show how the future prices are used to calculate the target surprise (Kansoy 2019; Hausman and Wongswan 2011) in the event study. The representation of the future rate ten minutes before the event will be:

$$ff_{t-10} = \frac{d}{D} r_0 + \frac{D-d}{D} E(r_1) + \eta_{t-10}$$

After the announcement is published, which is 20 minutes later, the future rate becomes:

$$ff_{t+20} = \frac{d}{D} r_0 + \frac{D-d}{D} r_1 + \eta_{t+20}$$

In these two equations, d represents days that have passed in the calendar month at the day of the announcement, D is the total days at the current month and stands for a risk premium. r_0 is the original federal fund rate before the change and r_1 represents the adjusted rate. We can see equations consist of two parts plus a risk premium part. The payoff of a one-month future contract is determined by the average of the effective federal fund rate. If a change of federal fund rate is expected to happen at time t in that month, the contract's payoff will be calculated using two different federal fund rates. Each part is weighted by the proportion of days that the corresponding rate has an effect on. The key point is that market participants will form an expected federal fund rate $E(r_1)$ before the actual rate reveals and use the expected value in the transactions.

I use these two equations to calculate the target surprise by applying the following calculation and assuming the risk premium stays the same.

$$ff_{t+20} - ff_{t-10} = \frac{(D-d)}{D} (r_1 - E(r_1))$$

By rearranging it, we get the target surprise:

$$r_1 - E(r_1) = \frac{D}{D-d} (ff_{t+20} - ff_{t-10})$$

The path surprise reflects the effect of news released by the central bank on people's expectations of future policy rates. The second factor is correlated with the variation of medium term and long-term interest rate. The proxy for the path surprise is the change of interest-rate-based eurodollar futures ($\triangle EF_t$). I label the variable as Path Surprise A. There is an alternative way of measuring the path surprise, that is to use the change of interest-rate-based eurodollar futures excluding the effect of target surprise. To extract the path surprise, I run the regression

$$\triangle EF_t = \alpha_0 + \alpha_1 TS_1 + \varepsilon_t$$

And the residual of the regression is the path surprise. The path surprise that is generated through this way is labelled as Path Surprise B

3.3 The basic analysis for FOMC minutes

Meeting minutes are published three weeks after the meeting. The policy information released from the statement is fully digested by the market and the minutes is unlikely to be the source of the target surprise. It is the 'soft information' contained in the content that has an impact on market participants' expectations. The variables include in the regression are mainly constructed in the content analysis and economic fundamentals indicators are added as control variables which will be discussed in the text analysis part.

3.4 Text analysis

Text data as a source of "soft information" is the main subject to be analysed in the second stage. By objectively and systematically quantifying certain features within words and sentences, researchers conduct content analysis to reveal how sentiment affects decision-makers and capture a relatively more objective picture of the object. Those qualitative characteristics of the textual data allow us to look into the price formation process from a different angle and add additional information to traditional asset pricing models (Kearney and Liu 2014). Generating those qualitative indicators from the raw text is the main challenge here.

I first extract content for both statements and minutes and construct variables in the document level.

$$Score_m^t = f_m^t$$

where f_m^t is the percentage of tone $m \in \{positive, negative, uncertain\}$ words in the document.

I implement a dictionary-based approach. There are two main approaches to quantifying central bank communication-manually-narrative approach and the dictionary-based word-count approach. Compared to the manual approach, the dictionary-based word count has the advantages of objectivity, continuity, and can be easily reproduced (Picault and Renault 2017). The programme maps the words in the documents to the pre-defined dictionary and labels them with the matching categories (positive words/negative words/words with uncertainty). Word lists and the weighting scheme of words in the word list are two important issues to be discussed in a dictionary-based approach (Kearney and Liu 2014). I use word lists developed by Loughran and McDonald (2011) (LM list). The reason is that this word list is constructed based on 10-Ks which is designed for analysing finance and business-related text materials. Another reason is its comprehensiveness compared to other word lists (Loughran and McDonald 2015). As for how the words are weighted, I follow Jegadeesh and Wu (2017)'s approach and directly use the frequency of the tonal words in the document as the term weighting scheme.

What do the tone change and uncertainty level change imply? Relative more negative words indicate the increasing possibility of tightening actions (ibid.). In Hansen and McMahon (2016)'s paper, the positive words are directly defined as expansion words and the negative words are called contraction words. Although the dictionary they used is different from the one I am using, they

emphasize the relationship between the tone and the potential actions.

As for the uncertainty level, Hansen and McMahon (2016) interpret it as a sign to the variance of future monetary policy surprise. Monetary policy uncertainty will affect the credit costs and real output (Husted, Rogers, and Sun 2020). The increase in uncertainty levels in the central bank communication makes it hard for the market to predict future policy path and leads to the increase in the credit costs.

Next section is about analysing the effect of different themes in the text, I apply the Latent Dirichlet Allocation (LDA) algorithm (Blei, Ng, and Jordan 2003) on FOMC minutes which is wildly used in the study of central bank communication. Statements are not suitable for a topic analysis because they are short articles (Jegadeesh and Wu 2017). The analysis of the statements will skip this step. The motivations for the analysis at the topic level are as followings. Jegadeesh and Wu (ibid.) mention that certain topics may contain more information than others. Boukus and Rosenberg (2006) also find that different themes have different impact on the change of the treasury yield. Priola, Molino, Tizzanini, et al. (2021) use the result of the LDA model and build topics weighted sentiment index for main central banks and found it has predictive power on the future policy rate. Luangaram, Wongwachara, et al. (2017) find one specific topic related to growth is informative in FOMC statement. The result of the analysis in topic levels can be instructive for policymakers. They can highlight specific topics to make most use of the central bank communication.

In general, the LDA model is a kind of generative model which is trying to explain the process of generating the data. In this case, how the text content of the announcement is formed. It simplifies reality by first ignoring the sequence of words and following the bag-of-words approach. Each document is generated word by word: the topic is decided first and word is drawn from the word distribution for each topic (Coelho, Richert, and Brucher 2018). What the algorithm does is to find topic distributions for each document and word distributions for each topic which maximized the probability of generating the bundles of words in the document analysed. As an unsupervised learning method with no pre-assigned labels, the LDA model can consistently generate topic distributions for each document and word distributions for each topic. Another advantage of the LDA model lies in its flexibility. Words in different topics are not mutually exclusive which increases the accuracy of capturing the content (Hansen, McMahon, and Prat 2018).

Scores that measure m tones which have three categories: positive, negative and uncertain for each paragraph d are constructed as follows:

$$Score^{t}_{d,n,m} = \widehat{\theta^{t}_{d,n}} F^{t}_{d,m}$$

in which t represents meeting time, n represents topic n in LDA model results, $\widehat{\theta}$ represents the estimated percentage of the topic and F is the count of tonal words.

We need to further calculate scores at the document level which can be repre-

sented as the aggregation of paragraph scores and multiplied by the weighting factor.

$$Score^t_{n,m} = \sum_{d=1}^{D_t} \widehat{\theta^t_{d,n}} F^t_{d,m}$$

where D_t is the total number of paragraphs of the meeting minute at time t and $f^t_{d,m}$ is $F^t_{d,m}$ divided by the total number of words in the paragraph d, which is the percentage of the tonal words in that paragraph.

3.5 The extended version of the model exploring the effect of the statements on asset prices

$$R_t = \alpha + \beta_1 TS_t + \beta_2 PS_t + \varepsilon_t$$

$$PS_t = \beta_3 Score^t_{nettone} + \beta_4 Score^t_{uncertaintylevel} + \varepsilon_t$$

The second regression is inspired by the critics made by Hansen and McMahon (2016) that there is no direct measurement of information conveyed by the central bank in GSS's design. The correlation between the tone changes and path surprise which is supposed to capture the news released to the public is examined in the second regression.

3.6 The extended version of the model exploring the effect of minutes on asset prices

The specification of the regression follows Boukus and Rosenberg (2006) and Jegadeesh and Wu (2017). In Boukus and Rosenberg's model, the uncertainty level is measured by the expected value of the volatility of Eurodollar futures. I replace it with the variables generated by the sentiment analysis which directly measures the uncertainty level. I also replace the term spread which is the variables indicate the status in the business cycle with other macroeconomic indicators used in Jegadeesh and Wu's model. The main idea, as Schmeling and Wagner (2019) summarise in their paper, is to examine whether tone change has explanatory power when all the control variables are held constant.

$$R_t = \alpha_0 + \alpha_1 Score^t_{nettone} + \alpha_2 Score^t_{uncertaintylevel} + \gamma M_t + \varepsilon_t$$

The explanation power of the scores for the whole document is examined first. M_t are the control variables: long-term interest rate, unemployment rate and a dummy variable indicate whether the economy is in recession.

The second regression is

$$R_t = \alpha_0 + \sum_{n=1}^{8} \alpha_n Score^t_{nettone} + \sum_{n=1}^{8} \alpha_{n+8} Score^t_{uncertaintylevel} + \gamma M_t + \varepsilon_t$$

in which net tone scores for each topic and uncertainty scores for each topic are added as independent variables.

4 Data Sources

The high frequency data is from the FirstRate data. All text data is directly downloaded from the federal reserve website by web scraping. As for the data used to construct the control variables for the analysis of the minutes, the dummy variable Recession and the monthly unemployment rate are from Federal Reserve Economic Data website. The long-term interest rate is from the Yahoo Finance.

5 Data description

5.1 Asset price data

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
EURUSD(%)	106	-0.00324	0.124	-0.298	0.317
CHFUSD(%)	106	0.0103	0.108	-0.246	0.250
JPYUSD(%)	106	0.00260	0.126	-0.340	0.301
SP500	106	-0.000119	0.00242	-0.00507	0.00505

Note: Asset price data is unavailable in 2020/3/15 and the observations are deleted

Figure 1: The summary statistics of the reaction of asset prices to FOMC statements

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
EURUSD (%)	106	-0.0115	0.179	-1.110	0.538
CHFUSD(%)	106	0.0221	0.164	-0.459	1.128
JPYUSD(%)	106	0.0110	0.172	-0.519	1.013
SP500(%)	106	-0.00644	0.266	-0.663	0.726

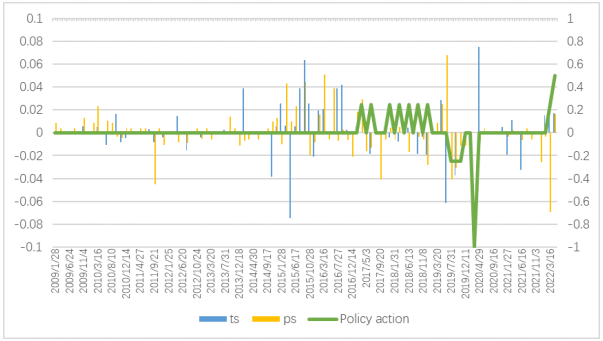
Note: CHFUSD exchange rate is unavailable in 2021/5/19 and this observation is deleted.

Figure 2: The summary statistics of the reaction of asset prices to FOMC minutes

Due to the availability of the high-frequency data from FirstRate data, the return of asset prices data from 2009.1-2022.5 is included. The variations of the return of asset prices in percentage at the observation time intervals (13:50-14:20) on the day when the statements and minutes are published are recorded and the total number of observations is 106. The exchange rates are adjusted

to the same format-the foreign currency to the US dollar. The summary tables of variables are listed above.

5.2 Target surprise and path surprise for statements



Note:Target surprise and Path surprise information in 2020/3/15 are missing. The policy action is the change of the upper limit of the target federal fund rate.

Figure 3: monetary policy surprise and Policy action for statements

Figure 3 presents the variation of target surprise, path surprise and policy action from 2009 Jan. to 2022 May. Two important time periods are included-the post global financial crisis and the Covid-19 pandemic period. To further verify whether the plot matches the situation in reality, I traced back the news at the turbulent time depicted in the figure 3. We can see from the plot that there is a relatively dramatic change in monetary policy surprise in mid-2015. At that time, the market was expecting an increase in policy rate but the Fed hesitated to act and leave the investors much uncertainty (Egan 2015).

Another point to discuss is that in Hausman and Wongswan (2011)'s paper, the target surprise shows a property of co-movement with the policy action in the time period from 1994 to 2005. However, the figure 3 does not show this property. A potential explanation is that the increase in the transparency of decision-making these years (Dincer and Eichengreen 2013) allows the market to predict the potential move of the Fed.

	Number of observations	Average word counts	Sample
Statements	117	749.2336	2009-2022
Minutes	117	8901.374	2009-2022

Figure 4: The Word Counts

I include 117 statements and minutes files as text materials to be analysed and the summary statistics are listed in the figure 4. The sample of 2021/5/19 is deleted in the analysis for FOMC minutes and the sample of 2020/3/15 is

deleted in the analysis for FOMC statements because the corresponding asset price data on these days is missing. The feature of the text data is that the length of the minutes is longer compared to the length of the statements.

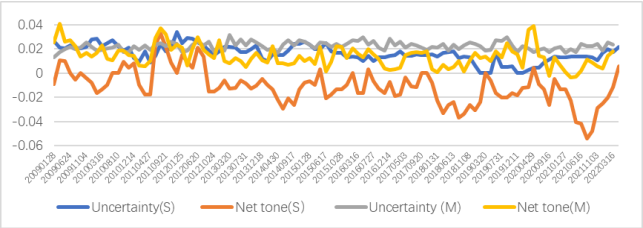


Figure 5: Sentiment Score for statements and minutes

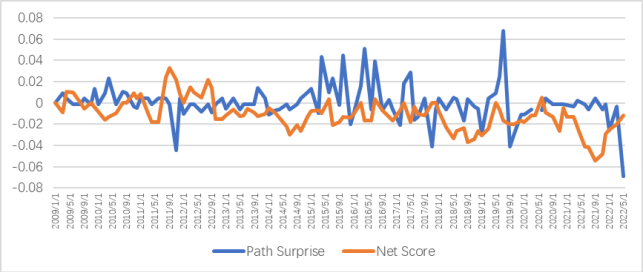


Figure 6: The change of net tone score and Path Surprise over time

Figure 5 shows the development of sentiment scores of statements and minutes. The net tone is the difference between negative scores and positive scores. If the score is greater than 0, it indicates there are more negative words in the statement than positive words. The most prominent feature is that the co-movement of sentiment scores of statements and minutes exists (the yellow line and the orange line). We can see there are increases in both pessimism sentiment and uncertainty levels during the mid-2011- 2012 period. During the covid-19 pandemic period, the trend of the net tone score goes upward in the early 2020s and moves downward in mid-2021. It starts to rise again after mid-2021. And the uncertainty level continues to grow during the pandemic period while there are small fluctuations.

If we add the movement of path surprise and put it with the change of net score of statements, as shown in Figure 6, we can see there is a negative correlation between path surprise and net score.

The analysis for the minutes at the topic level is more complex than the analysis for statements. There are a few more steps to do to transform raw text materials to the input of topic modelling. The first step is to remove all special characters in HTML and redundant spaces. The second step is to clean the text is to delete all stop words like is ,are, the which contain no meaning. Punctuation is removed as well. In the process to group words in different forms but with

the same meanings, I choose lemmatization which returns the original form of words. Because compare to stemming which removes the inflectional endings of words, lemmatization has better performance in precision (Balakrishnan and Lloyd-Yemoh 2014). This is the reason why the words in the word clouds are all complete and in their original form. Some researchers apply both stemming and lemmatization in their word lists.

A dictionary used to store word ID is constructed using a training corpus. The purpose of this step is to build a list that records the frequency of each unique word in its ID number form. Again, the sequence of words here is ignored and this is the bag-of-words approach.

The final step is to choose the parameters for LDA models. I choose the best combination of parameters based on the coherence value of corresponding LDA models.

There are eight topics in total and each topic is assigned a theme according to its keywords. The keywords in the first topic are *security, purchase, agency, policy, asset, mortgage, back, holding, balance, market* , which are related to market. The second topic contains *inflation, price, labor, measure, rate, remain.* I assign Inflation as the theme of this topic. *Market, period, intermeete, loan, yield, credit* are in the third topic and they are also related to market. The fourth topic is related to monetary policy which includes *inflation, policy, rate, economic, federal, fund. Consumption is the theme for the fifth topic and its keywords are quarter, spending, business, increase, real, consumer, decline, rise, production, good.* The sixth topic is about discussions about the foreign economy, which includes words like *rate, operation, foreign, market, transaction, currency, vote, swap, direct, facility.* Keywords like *participant, economic, note, risk, financial, economy, outlook, activity, continue, market* reflect this topic is about financial stability. The last topic is about economic growth, as indicated by the words *growth, year, staff, rate, real, quarter, continue, economic, gdp, projection.*

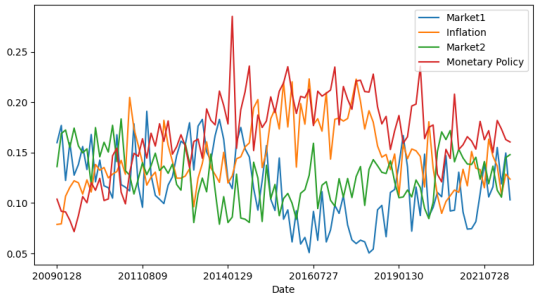


Figure 7: Topic (1-4) proportions over time

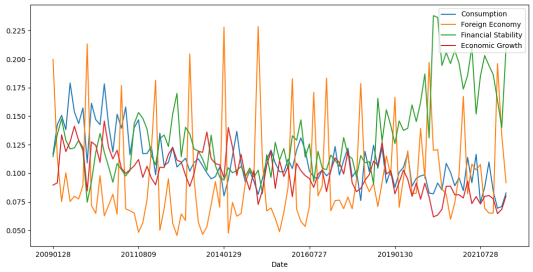


Figure 8: Topic (5-8) proportions over time

Figures 7,8 show the topic proportions change over time. I compare the result distribution with the LDA topic modelling results of Jegadeesh and Wu (2017), Benchimol, Kazinnik, and Saadon (2021). Models with different parameters generate different results and economists have different interpretations of the keywords of each topic. The key feature of JW's result is that there has been less attention to economic growth since 2004. BWS's model shares a similar pattern. It can be found in my model as well. The "Inflation" topic is becoming more and more important, which can be found in all three model's results. Although there are fluctuations. In BKS's model, the topic "Financial stability" attracts more attention after the 2008 financial crisis. The proportion of the corresponding topic "market" in JW's model also increases after 2008. The proportion of the "Financial stability" topic in my model is relatively stable during 2009-2018. The potential explanation can be the absence of data before 2009. There is no comparison. The percentage of "foreign economy" is the most turbulent part of my model which can not be found in JW's model.

6 Regression results

6.1 Statement

We first consider the original Hausman and Wongswan (2011)'s model using the updated data and the regression results are shown below. Two different measurements of the path surprise are applied and the results are listed separately. The R-squared of the model using Path Surprise B is larger than the one using Path Surprise A. The dependent variables have little explanation of the variation of the exchange rate after the announcement and the results are mostly statistically insignificant. The CHF to USD exchange rate reacts differently from the other two exchange rates. The directions of the impact of target surprise on the stock market are different in models using different proxies for path surprise. For the model using Path Surprise A, a hypothetical increase of 1% in target surprise is associated with a 0.825% decline in SP 500 return. However, the result of the model using Path Surprise B does not match the previous empirical analysis results. The increment in target surprise is correlated with a negative shock to the stock market in the precious research.

$$R_t = \alpha + \beta_1 TS_t + \beta_2 PS_t + \varepsilon_t$$

VARIABLES	(1) SP_500(%)	(2) Eurusd(%)	(3) Jpyusd(%)	(4) Chfusd(%)
Target Surprise	-0.825	-0.805	-0.334	0.831
	(2.256)	(0.808)	(0.723)	(0.687)
Path Surprise A	9.365***	0.386	0.259	-0.998
	(2.064)	(0.739)	(0.661)	(0.629)
Constant	0.0429	0.0107	0.00306	-0.00808
	(0.0324)	(0.0116)	(0.0104)	(0.00987)
Observations	106	106	106	106
R-squared	0.170	0.010	0.003	0.031

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 9: The regression of market returns on Target Surprise and Path Surprise A(Statements)

VARIABLES	(1) SP_500	(2) Eurusd	(3) Jpyusd	(4) Chfusd
Target Surprise	0.00643	-0.140	-0.214	1.985***
	(0.0173)	(1.392)	(0.728)	(0.667)
Path Surprise B	0.0849***	0.767	0.411	-0.992
	(0.0182)	(1.460)	(0.764)	(0.699)
Constant	0.000522	-0.000838	-0.000800	-0.0152
	(0.000322)	(0.0258)	(0.0135)	(0.0124)
Observations	106	106	106	106
R-squared	0.176	0.003	0.004	0.095

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 10: The regression of market returns on Target Surprise and Path Surprise B(Statements)

To investigate whether the tones and uncertainty level of the statements are correlated with the path surprise, I run the second regression. We can see the models using two different measurements generate similar results. The increment of negative words in the statements relative to positive words is correlated with a larger path surprise. However, the uncertainty level is negatively related to the magnitude of the path surprise.

$$PS_t = \beta_3 Score^t_{nettone} + \beta_4 Score^t_{uncertaintylevel} + \varepsilon_t$$

VARIABLES	(1) Path Surprise A	(2) Path Surprise B
Uncertainty Level	-0.000926	-0.000905
	(0.000812)	(0.000794)
Net Tone	0.0315	0.0153
	(0.107)	(0.105)
Constant	0.00140	-1.77e-05
	(0.00195)	(0.00190)
Observations	106	106
R-squared	0.014	0.013

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 11: The regression of path surprise on sentiment scores(statement)

6.2 Minutes

I first examine the relationship between the sentiment change at the document level and the change of return in the stock market and foreign currency markets. The percentage of the negative words minus the per cent positive words in the statement is positively correlated with the SP 500 return and the exchange rate of EUR to USD. However, the correlation of the tone change with the reaction of the exchange rate of JPY and CHF is reversed.

VARIABLES	(1) SP_500 textual analysis	(2) EURUSD textual analysis	(3) JPYUSD textual analysis	(4) CHFUSD textual analysis
Net tone	3.514	4.502*	-4.370*	-1.677
	(3.632)	(2.425)	(2.335)	(2.230)
Uncertainty	1.409	-0.658	0.785	-4.274
	(8.534)	(5.699)	(5.488)	(5.241)
Interest rate	0.0205	0.00989	-0.00873	-0.0201
	(0.0366)	(0.0245)	(0.0235)	(0.0225)
Unemployment rate	-0.00641	-0.00479	0.00520	0.000897
	(0.0133)	(0.00889)	(0.00856)	(0.00817)
Recession	0.0767	-0.0639	0.0572	-0.0841
	(0.158)	(0.106)	(0.102)	(0.0973)
Constant	-0.0953	-0.0487	0.0385	0.184
	(0.233)	(0.156)	(0.150)	(0.143)
Observations	106	106	106	106
R-squared	0.020	0.036	0.036	0.034

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 12: The regression of return of asset prices on sentiment scores of minutes(document level)

I further use the result of the LDA model to investigate the effect of different topics in the minutes on the market. NT 0-7 represent the net tone of each topic and UN 0-7 are uncertainty level of each topic. Topics 0-7 are market, inflation, another market-related topic, monetary policy, consumption, foreign economy, financial stability and economic growth respectively. We can see the directions of the impact of the tone change and uncertainty level of different themes are not consistent. This feature can also be found in Boukus and Rosenberg (2006) 's result. The increasing pessimism conveyed in three topics is related to a decline in the return of the SP 500. Those topics are monetary policy, consumption and

financial stability. As for the uncertainty level, the increase of the uncertainty level in four topics is related to a decrease in the return of the SP 500. Those topics are market, inflation, consumption and foreign economy.

As for the exchange rate, the coefficients of the net tone in inflation, market and economic growth topics are negative, which means an increase in pessimism in those three topics is related to a depreciation of the USD. There are five topics out of seven topics which have negative estimated parameters in the regression using the exchange rate JPY to USD as the dependent variable. The net tone of the Inflation and financial stability topics is negatively correlated with the exchange rate of CHF to USD. In summary, we can not get a clear conclusion of the direction of the impact of the tone of each topic. This applies to the result of the uncertainty level as well.

VARIABLES	(1) SP_500	(2) Eurusd	(3) Jpyusd	(4) Chfusd
NT0	1.343*	0.184	-0.221	0.257
	(0.704)	(0.532)	(0.515)	(0.489)
NT1	1.083***	0.143	-0.157	-0.460*
	(0.369)	(0.279)	(0.270)	(0.256)
NT2	0.226	-0.479*	0.423*	0.338
	(0.327)	(0.247)	(0.239)	(0.227)
NT3	-1.370***	-0.135	0.137	0.220
	(0.379)	(0.287)	(0.278)	(0.263)
NT4	-0.900**	0.131	-0.154	0.0177
	(0.343)	(0.259)	(0.251)	(0.238)
NT5	1.377***	0.676*	-0.612*	0.00732
	(0.499)	(0.377)	(0.365)	(0.346)
NT6	-0.244	0.440	-0.391	-0.519*
	(0.387)	(0.292)	(0.283)	(0.268)
NT7	0.752	-0.0624	0.107	0.290
	(0.558)	(0.422)	(0.408)	(0.387)
UN0	-0.449	0.0498	0.0284	-0.655
	(0.894)	(0.676)	(0.654)	(0.621)
UN1	-0.555	0.262	-0.217	-0.224
	(0.814)	(0.616)	(0.596)	(0.566)
UN2	0.943	0.621	-0.565	-0.275
	(0.636)	(0.481)	(0.465)	(0.442)
UN3	0.990*	-0.238	0.231	-0.186
	(0.529)	(0.400)	(0.387)	(0.367)
UN4	-1.033	-0.331	0.333	-0.542
	(0.825)	(0.624)	(0.604)	(0.573)
UN5	-0.822*	-0.00498	-0.0131	0.319
	(0.491)	(0.372)	(0.360)	(0.341)
UN6	0.329	-0.272	0.242	0.478
	(0.449)	(0.340)	(0.329)	(0.312)
UN7	1.148	0.00905	-0.163	0.353
	(1.154)	(0.872)	(0.844)	(0.801)
Interest rate	0.0523	0.0141	-0.0137	-0.0356
	(0.0492)	(0.0372)	(0.0360)	(0.0342)
Unemployment rate	0.0172	-0.0135	0.0123	0.00678
	(0.0199)	(0.0150)	(0.0146)	(0.0138)
Recession	0.0390	-0.144	0.135	-0.0824
	(0.159)	(0.120)	(0.116)	(0.110)
Constant	-0.727**	0.0565	-0.0394	0.177
	(0.313)	(0.237)	(0.229)	(0.217)
Observations	106	106	106	106
R-squared	0.322	0.144	0.136	0.145

This table shows the result of the regression of return of asset prices on sentiment scores of minutes(at the topic level)

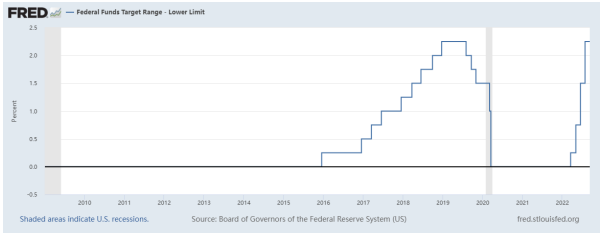
Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.3 the Zero lower bound period

In this part, I use the sample period when the overnight rate hits 0 to investigate whether the effect of tone change and uncertain levels in central bank communication on the financial market during the zero lower bound period has some new features. As shown in the figure 13, the federal fund rate is near 0 during 2009-2015.12 and 2020.3-2022.3. After excluding the missing data sample 2020/3/15, I include 71 samples in the analysis of the statements. In the first regression, most estimated parameters are larger compared to the result using whole sample periods while there are few exceptions.

On the second regression, models using only zero lower bound sample periods have higher R squared value compared to the models that cover the whole sample periods. Another point worth noting is that the direction of the impact of the net tone of the model using zero lower bound periods switches to negative, which means the increase of the relative percentage of the negative words is correlated with a smaller path surprise. As for the uncertainty level, the release of statements with a higher percentage of uncertain words is related to a higher path surprise to the market.



Source:https://fred.stlouisfed.org

Figure 13: Federal Fund rate from 2009-2022

VARIABLES	(1) SP_500	(2) Eurusd	(3) Jpyusd	(4) Chfusd
Target Surprise	-1.005	-1.092	-0.0683	0.570
	(2.910)	(1.220)	(1.069)	(1.009)
Path Surprise A	12.00***	0.412	1.929	-2.121
	(4.011)	(1.681)	(1.473)	(1.391)
Constant	0.0436	0.00751	-0.00112	-0.00766
	(0.0390)	(0.0164)	(0.0143)	(0.0135)
Observations	71	71	71	71
R-squared	0.116	0.012	0.025	0.035

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 14: The regression of market returns on Target Surprise and Path Surprise A(statement ZLB)

VARIABLES	(1) SP_500	(2) Eurusd	(3) Jpyusd	(4) Chfusd
Target Surprise	0.927	-0.987	0.262	0.206
	(2.897)	(1.209)	(1.057)	(0.997)
Path Surprise B	11.57***	1.037	2.183	-2.422*
	(3.949)	(1.647)	(1.441)	(1.358)
Constant	0.0545	0.00774	0.000560	-0.00949
	(0.0389)	(0.0162)	(0.0142)	(0.0134)
Observations	71	71	71	71
R-squared	0.112	0.017	0.033	0.046

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 15: The regression of market returns on Target Surprise and Path Surprise B(statement ZLB)

VARIABLES	(1) Path Surprise A	(2) Path Surprise B
Uncertainty	0.252	0.352*
	(0.206)	(0.207)
Net tone	-0.0578	-0.0748
	(0.0777)	(0.0778)
Constant	-0.00393	-0.00720
	(0.00433)	(0.00433)
Observations	71	71
R-squared	0.023	0.043

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 16: The regression of path surprise on sentiment scores(statement ZLB)

The next section is about analysing the market response to the release of minutes during the zero lower bound period. The first regression is using net tone and uncertainty at the document level. Compared to the results using the whole sample periods, there is a larger impact of the sentiment change in the Fed communication during the zero lower bound period. A larger R squared value compared to the previous model indicates the explanation power is also higher.

VARIABLES	(1) SP_500 textual analysis	(2) EURUSD textual analysis	(3) JPYUSD textual analysis	(4) CHFUSD textu analysis
Net Tone	5.051	5.396	-5.314	-1.663
	(4.784)	(3.264)	(3.189)	(3.155)
Uncertainty	9.282	4.804	-3.863	-7.330
	(11.70)	(7.984)	(7.799)	(7.717)
Interest rate	0.0198	0.0153	-0.0138	-0.0230
	(0.0443)	(0.0303)	(0.0296)	(0.0292)
Unemployment rate	-0.0150	-0.00701	0.00772	-0.00448
	(0.0218)	(0.0149)	(0.0145)	(0.0144)
Recession	0.0996	-0.0506	0.0475	-0.0883
	(0.180)	(0.123)	(0.120)	(0.119)
Constant	-0.217	-0.178	0.147	0.305
	(0.322)	(0.220)	(0.215)	(0.213)
Observations	71	71	71	71
R-squared	0.046	0.062	0.058	0.054

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 17: The regression of return of asset prices on sentiment scores of minutes(document level ZLB)

The next part is the discussion about the impact of minutes at the topic level. Similar results of relatively high R squared value can also be found in the regression results. The influence of sentiment changes on each topic becomes larger during the zero lower bound period relative to the complete sample periods. Another finding is that parameters of uncertainty level change in inflation, market, monetary policy and foreign economy topics are statistically significant.

VARIABLES	(1) SP_500	(2) Eurusd	(3) Jpyusd	(4) Chfusd
NT0	1.931*	0.762	-0.748	0.149
	(0.972)	(0.857)	(0.843)	(0.819)
NT1	1.294**	0.370	-0.372	-0.765*
	(0.507)	(0.447)	(0.439)	(0.427)
NT2	0.0344	-0.722*	0.628*	0.564
	(0.423)	(0.373)	(0.367)	(0.356)
NT3	-1.999***	-0.120	0.135	0.220
	(0.547)	(0.482)	(0.474)	(0.461)
NT4	-0.779*	-0.0164	-0.0331	0.0283
	(0.427)	(0.377)	(0.370)	(0.360)
NT5	1.538**	0.537	-0.499	0.0874
	(0.610)	(0.538)	(0.529)	(0.514)
NT6	-0.367	0.417	-0.360	-0.456
	(0.455)	(0.401)	(0.394)	(0.383)
NT7	0.853	-0.127	0.193	0.480
	(0.673)	(0.594)	(0.584)	(0.567)
UN0	-0.333	-0.270	0.386	-0.609
	(1.340)	(1.181)	(1.161)	(1.129)
UN1	-2.277**	0.624	-0.569	-0.489
	(1.021)	(0.901)	(0.885)	(0.861)
UN2	2.246**	1.002	-0.914	-0.585
	(0.866)	(0.764)	(0.751)	(0.730)
UN3	2.912***	-0.252	0.259	-0.488
	(0.832)	(0.734)	(0.721)	(0.701)
UN4	0.615	-0.664	0.759	-1.067
	(1.327)	(1.170)	(1.150)	(1.118)
UN5	-2.470***	-0.167	0.104	0.939
	(0.722)	(0.637)	(0.626)	(0.609)
UN6	-0.251	0.0454	-0.0746	0.545
	(0.602)	(0.531)	(0.522)	(0.507)
UN7	0.317	0.240	-0.437	0.970
	(1.450)	(1.279)	(1.257)	(1.222)
Interest rate	0.102	0.0287	-0.0308	-0.0276
	(0.0709)	(0.0625)	(0.0615)	(0.0597)
Unemployment rate	0.0193	0.00179	-0.00200	-0.00744
	(0.0306)	(0.0270)	(0.0265)	(0.0258)
Recession	0.00878	-0.101	0.0971	-0.138
	(0.165)	(0.145)	(0.143)	(0.139)
Constant	-1.049**	-0.186	0.189	0.310
	(0.465)	(0.410)	(0.403)	(0.392)
Observations	71	71	71	71
R-squared	0.512	0.199	0.186	0.211

This table shows the result of the regression of return of asset prices on sentiment scores of minutes(Topic level at ZLB period)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.4 Discussion of the findings

In general, the tone change of the FOMC communication has limited explanation power in the variation of the stock market and foreign currency market. It matches the results of Hansen and McMahon (2016)'s paper. They found communication content in the FOMC statement has no significant effect on real economic indicators.

The second finding is that there are more statistically significant estimated parameters in the regression using the return of the SP 500 as the dependent variable. This implies participants in the stock market find the Fed communication informative. However, the impact of communication on the exchange rates is limited in this model, which is contrary to the finding of Tadle (2022). He finds that there is a statistically significant correlation between a hawkish tone in meeting minutes and USD appreciation again JYP and CHF. This relationship applies to EUR as well, although it is not statistically significant.

The third finding is that content analysis indicators have a better explanation of the variation of the return of the stock markets and foreign currency markets during the zero lower bound period compared to the result using the whole sample period. This finding provides supporting evidence that communication becomes particularly important during the zero lower bound periods.

The analysis of minutes at the topic level reveals that the net tone changes of topics market, inflation, monetary policy, consumption, foreign economy and the uncertainty level of topics monetary policy and foreign economy have a statistically significant impact on the stock market. The relationship is not very clear in the foreign currency markets.

7 Robustness Check

I use the daily data of the asset prices to check the robustness of the results. The regression results are listed in the appendix. The first finding can be found in the results using alternative data of market reaction. The number of parameters that are statistically significant is small. The second finding is weakened. As for the third finding, it still applies to the results of the analysis of minutes. It does not hold for the results of the analysis of statements. At the topics level, the net tone changes of topics inflation, monetary policy and the uncertainty level of the topics monetary policy and foreign economy still have statistically significant impact on the magnitude of the stock market reaction to the release of the FOMC meeting minutes.

8 Limitations

The first limitation is in choosing parameters for the LDA model. The parameters are chosen based on comparative statics. Optimal parameters are chosen separately and models for comparisons are generated by varying one parameter while holding other parameters constant. A better solution will be grid search

which requires many computing resources. It will test the combination of parameters exhaustively.

Another point is that there would be less distraction if all introductory and administrative-related parts of the FOMC documents are deleted. Those parts contain the names of people who attend the meeting and the voting results.

There is another way to improve the result. There can be more subdivided parts of the text content. Monetary policy communication can be divided into two parts-the economic outlooks and forward guidance (Hansen and McMahon 2016). The forward guidance can further be grouped into two different categories. Delphic forward guidance reduces market uncertainty by offering better information about macroeconomic indicators, which refers to the fed information effect. The second type-Odyssean forward guidance is about making promises. The key difference between the two forward guidance is whether the central bank makes the commitment (Campbell et al. 2012). The results would be better if scores are generated for each subdivided part of the content of the communication.

In the section exploring the effect of minutes on the market, the result can be improved by extracting the incremental informativeness of minutes relative to statements. After all, it is the new information in the minutes which statements do not have that brings the impact.

The response of the stock market and the foreign currency markets are examined in this paper. The yield of treasury bills which reflects the interest rate is an important economic indicator. The inclusion of this marker would offer a more comprehensive result of how monetary policy communication affects the financial market compared to the current model

In the process of interpreting topic keywords and assigning a theme to each topic, it would be more precise and convincing to have verifications. In Boukus and Rosenberg (2006)'s paper, the correlation between topics and macroeconomic indicators is used to show the validity of the model. Some researchers like Jegadeesh and Wu (2017) hire human readers to identify the theme of the original paragraph of the minutes.

9 Conclusion

Based on the existing model, this paper employs a computational linguistic approach to investigate the effect of the FOMC statements and minutes on the stock markets and the foreign currency market by quantifying the sentiment change of the FOMC communication. The results show that the net tone changes and the uncertainty level have limited explanatory power. Also, the stock market reacts more to the net tone and uncertainty level of the FOMC policy documents compared to the foreign currency market. The third finding is that communication becomes particularly important during the zero lower bound period. The fourth finding is that there is a statistically significant correlation between the net tone of topics market, inflation, monetary policy,

consumption, foreign economy, the uncertainty level of topics monetary policy and foreign economy of the FOMC meeting minutes and the reaction of the stock market to the FOMC minutes.

References

Apel, Mikael and Marianna Blix Grimaldi (2014). "How informative are central bank minutes?" In: *Review of Economics* 65.1, pp. 53–76.

Balakrishnan, Vimala and Ethel Lloyd-Yemoh (2014). "Stemming and lemmatization: A comparison of retrieval performances". In.

Benchimol, Jonathan, Sophia Kazinnik, and Yossi Saadon (2021). "Federal reserve communication and the COVID-19 pandemic". In: *Covid Economics* 79, pp. 218–256.

Binder, John (1998). "The event study methodology since 1969". In: *Review of quantitative Finance and Accounting* 11.2, pp. 111–137.

Blei, David M, Andrew Y Ng, and Michael I Jordan (2003). "Latent dirichlet allocation". In: *Journal of machine Learning research* 3.Jan, pp. 993–1022.

Blinder, Alan S et al. (2008). "Central bank communication and monetary policy: A survey of theory and evidence". In: *Journal of economic literature* 46.4, pp. 910–45.

Boukus, Ellyn and Joshua V Rosenberg (2006). "The information content of FOMC minutes". In: *Available at SSRN 922312*.

Campbell, Jeffrey R et al. (2012). "Macroeconomic effects of federal reserve forward guidance [with comments and discussion]". In: *Brookings papers on economic activity*, pp. 1–80.

Cieslak, Anna and Andreas Schrimpf (2019). "Non-monetary news in central bank communication". In: *Journal of International Economics* 118, pp. 293–315.

Coelho, Luis Pedro, Willi Richert, and Matthieu Brucher (2018). *Building Machine Learning Systems with Python: Explore machine learning and deep learning techniques for building intelligent systems using scikit-learn and TensorFlow*. Packt Publishing Ltd.

Danker, Deborah J and Matthew M Luecke (2005). "Background on FOMC meeting minutes". In: *Fed. Res. Bull.* 91, p. 175.

Dincer, N Nergiz and Barry Eichengreen (2013). "Central bank transparency and independence: updates and new measures". In.

Egan (2015). *Wall Street struggles to decode Fed decision*. URL: `https://money.cnn.com/2015/09/17/investing/fed-rate-hike-stocks-markets/index.html`.

Fawley, Brett W, Christopher J Neely, et al. (2014). "The evolution of Federal Reserve policy and the impact of monetary policy surprises on asset prices". In: *Federal Reserve Bank of St. Louis Review* 96.1, pp. 73–109.

Gürkaynak, Refet S, Brian P Sack, and Eric T Swanson (2004). "Do actions speak louder than words? The response of asset prices to monetary policy actions and statements". In: *SSRN Electronic Journal*.

Hamilton, James D (2009). "Daily changes in fed funds futures prices". In: *Journal of Money, credit and Banking* 41.4, pp. 567–582.

Hansen, Stephen and Michael McMahon (2016). "Shocking language: Understanding the macroeconomic effects of central bank communication". In: *Journal of International Economics* 99, S114–S133.

Hansen, Stephen, Michael McMahon, and Andrea Prat (2018). "Transparency and deliberation within the FOMC: a computational linguistics approach". In: *The Quarterly Journal of Economics* 133.2, pp. 801–870.

Hausman, Joshua and Jon Wongswan (2011). "Global asset prices and FOMC announcements". In: *Journal of International Money and Finance* 30.3, pp. 547–571.

Husted, Lucas, John Rogers, and Bo Sun (2020). "Monetary policy uncertainty". In: *Journal of Monetary Economics* 115, pp. 20–36.

Jansen, David-Jan and Jakob De Haan (2007). "The importance of being vigilant: Has ecb communication influenced euro area inflation expectations?" In: *Available at SSRN 1027576*.

Jegadeesh, Narasimhan and Di Wu (2017). "Deciphering Fedspeak: The information content of FOMC meetings". In: *SSRN, https://ssrn. com/abstract* 2939937.

Kansoy, Fatih (2019). "FOMC minutes: as a source of monetary policy surprise". In: *manuscript, University of Nottingham*.

Kearney, Colm and Sha Liu (2014). "Textual sentiment in finance: A survey of methods and models". In: *International Review of Financial Analysis* 33, pp. 171–185.

Kuttner, Kenneth N (2001). "Monetary policy surprises and interest rates: Evidence from the Fed funds futures market". In: *Journal of monetary economics* 47.3, pp. 523–544.

Loughran, Tim and Bill McDonald (2011). "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks". In: *The Journal of finance* 66.1, pp. 35–65.

— (2015). "The use of word lists in textual analysis". In: *Journal of Behavioral Finance* 16.1, pp. 1–11.

Luangaram, Pongsak, Warapong Wongwachara, et al. (2017). "More than words: a textual analysis of monetary policy communication". In: *PIER Discussion Papers* 54.

Nakamura, Emi and Jón Steinsson (2018). "High-frequency identification of monetary non-neutrality: the information effect". In: *The Quarterly Journal of Economics* 133.3, pp. 1283–1330.

Nechio, Fernanda, Daniel J Wilson, et al. (2016). "How important is information from FOMC minutes?" In: *FRBSF Economic Letter* 37.

Picault, Matthieu and Thomas Renault (2017). "Words are not all created equal: A new measure of ECB communication". In: *Journal of International Money and Finance* 79, pp. 136–156.

Pinchetti, Marco and Andrzej Szczepaniak (2021). "Global spillovers of the fed information effect". In.

Priola, Maria Paola, Annalisa Molino, Giacomo Tizzanini, et al. (2021). "The informative value of central banks talks: a topic model application to sentiment analysis". In.

Romer, Christina D and David H Romer (2000). "Federal Reserve information and the behavior of interest rates". In: *American economic review* 90.3, pp. 429–457.

Rosa, Carlo (2011). "The high-frequency response of exchange rates to monetary policy actions and statements". In: *Journal of Banking & Finance* 35.2, pp. 478–489.

— (2013). "The financial market effect of FOMC minutes". In: *Available at SSRN 2378398*.

Rudebusch, Glenn et al. (2018). "A review of the fed's unconventional monetary policy". In: *FRBSF Economic Letter* 27, pp. 1–5.

Schmeling, Maik and Christian Wagner (2019). "Does central bank tone move asset prices?" In: *Available at SSRN 2629978*.

Tadle, Raul Cruz (2022). "FOMC minutes sentiments and their impact on financial markets". In: *Journal of economics and business* 118, p. 106021.

Appendix

(1) The release date of minutes and statements

The release date of statements	The release date of minutes
2009/1/28	2009/2/18
2009/3/17	2009/4/8
2009/4/29	2009/5/20
2009/6/24	2009/7/15
2009/8/11	2009/9/2
2009/9/22	2009/10/14
2009/11/4	2009/11/24
2009/12/15	2010/1/6
2010/1/27	2010/2/17
2010/3/16	2010/4/6
2010/4/28	2010/5/19
2010/6/23	2010/7/14
2010/8/10	2010/8/31
2010/9/21	2010/10/12
2010/11/3	2010/11/23
2010/12/14	2011/1/4
2011/1/26	2011/2/16
2011/3/15	2011/4/5
2011/4/27	2011/5/18
2011/6/22	2011/7/12
2011/8/9	2011/8/30
2011/9/21	2011/10/12
2011/11/2	2011/11/22
2011/12/13	2012/1/3
2012/1/25	2012/2/15
2012/3/13	2012/4/3
2012/4/25	2012/5/16
2012/6/20	2012/7/11
2012/8/1	2012/8/22
2012/9/13	2012/10/4
2012/10/24	2012/11/14
2012/12/12	2013/1/3
2013/1/30	2013/2/20
2013/3/20	2013/4/10
2013/5/1	2013/5/22
2013/6/19	2013/7/10
2013/7/31	2013/8/21
2013/9/18	2013/10/9
2013/10/30	2013/11/20
2013/12/18	2014/1/8
2014/1/29	2014/2/19

2014/3/19	2014/4/9
2014/4/30	2014/5/21
2014/6/18	2014/7/9
2014/7/30	2014/8/20
2014/9/17	2014/10/8
2014/10/29	2014/11/19
2014/12/17	2015/1/7
2015/1/28	2015/2/18
2015/3/18	2015/4/8
2015/4/29	2015/5/20
2015/6/17	2015/7/8
2015/7/29	2015/8/19
2015/9/17	2015/10/8
2015/10/28	2015/11/18
2015/12/16	2016/1/6
2016/1/27	2016/2/17
2016/3/16	2016/4/6
2016/4/27	2016/5/18
2016/6/15	2016/7/6
2016/7/27	2016/8/17
2016/9/21	2016/10/12
2016/11/2	2016/11/23
2016/12/14	2017/1/4
2017/2/1	2017/2/22
2017/3/15	2017/4/5
2017/5/3	2017/5/24
2017/6/14	2017/7/5
2017/7/26	2017/8/16
2017/9/20	2017/10/11
2017/11/1	2017/11/22
2017/12/13	2018/1/3
2018/1/31	2018/2/21
2018/3/21	2018/4/11
2018/5/2	2018/5/23
2018/6/13	2018/7/5
2018/8/1	2018/8/22
2018/9/26	2018/10/17
2018/11/8	2018/11/29
2018/12/19	2019/1/9
2019/1/30	2019/2/20
2019/3/20	2019/4/10
2019/5/1	2019/5/22
2019/6/19	2019/7/10
2019/7/31	2019/8/21
2019/9/18	2019/10/9
2019/10/30	2019/11/20
2019/12/11	2020/1/3

2020/1/29	2020/2/19
2020/3/15 (missing)	2020/4/8
2020/4/29	2020/5/20
2020/6/10	2020/7/1
2020/7/29	2020/8/19
2020/9/16	2020/10/7
2020/11/5	2020/11/25
2020/12/16	2021/1/6
2021/1/27	2021/2/17
2021/3/17	2021/4/7
2021/4/28	2021/5/19 (missing)
2021/6/16	2021/7/7
2021/7/28	2021/8/18
2021/9/22	2021/10/13
2021/11/3	2021/11/24
2021/12/15	2022/1/5
2022/1/26	2022/2/16
2022/3/16	2022/4/6
2022/5/4	2022/5/25

Regression results of robustness checking

Statement: The daily data is from Yahoo Finance. Data in 2020/3/15 is missing and the total number of observations is 106.

	(1)	(2)	(3)	(4)
VARIABLES	SP_500	Chfusd	Eurusd	Jpyusd
Target Surprise	10.20	-1.037	-1.617	3.246
	(6.960)	(3.410)	(3.023)	(2.971)
Path Surprise A	5.782	1.399	-1.972	1.006
	(6.369)	(3.121)	(2.766)	(2.718)
Constant	0.247**	0.0630	0.0774*	-0.0219
	(0.1000)	(0.0490)	(0.0434)	(0.0427)
Observations	106	106	106	106
R-squared	0.035	0.002	0.010	0.015

(2) The regression result of market returns on Target Surprise and Path Surprise A(Statement_Daily Data)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) SP_500	(2) Chfusd	(3) Eurusd	(4) Jpyusd
Target Surprise	11.08	-0.841	-1.926	3.393
	(6.814)	(3.337)	(2.960)	(2.907)
Path Surprise B	5.330	1.436	-1.716	0.971
	(6.381)	(3.125)	(2.772)	(2.722)
Constant	0.253**	0.0645	0.0754*	-0.0209
	(0.0999)	(0.0489)	(0.0434)	(0.0426)
Observations	106	106	106	106
R-squared	0.034	0.002	0.009	0.015

(3) The regression result of market returns on Target Surprise and Path Surprise B(statement_ Daily Data)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

During the Zero lower bound period

VARIABLES	(1) SP_500	(2) Eurusd	(3) Jpyusd	(4) Chfusd
Target Surprise	13.21	-2.923	1.381	-1.240
	(10.04)	(4.398)	(4.297)	(5.091)
Path Surprise A	0.499	4.212	0.596	9.778
	(13.99)	(6.125)	(5.985)	(7.090)
Constant	0.384***	0.0896	-0.0235	0.0850
	(0.135)	(0.0592)	(0.0579)	(0.0686)
Observations	72	72	72	72
R-squared	0.025	0.011	0.002	0.027

(4) The regression result of market returns on Target Surprise and Path Surprise A (Statement_ZLB_Daily Data)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) SP_500	(2) Eurusd	(3) Jpyusd	(4) Chfusd
Target Surprise	13.22	-2.216	1.466	0.325
	(9.923)	(4.341)	(4.247)	(5.035)
Path Surprise B	-1.049	4.738	0.332	9.281
	(13.73)	(6.008)	(5.877)	(6.968)
Constant	0.384***	0.0934	-0.0229	0.0939
	(0.135)	(0.0589)	(0.0576)	(0.0683)
Observations	72	72	72	72
R-squared	0.025	0.013	0.002	0.025

(5) The regression result of market returns on Target Surprise and Path Surprise B (Statement_ZLB_Daily Data)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Minutes: Data for day 2010/10/11,2012/1/2, 2013/10/8, 2017/7/4, 2018/7/4, 2019/5/22 and 2021/5/19 is missing.

VARIABLES	(1) SP_500	(2) EURUSD	(3) JPYUSD	(4) CHFUSD
Net tone(w)	19.57*	-6.934	0.108	-7.829
	(11.30)	(7.406)	(7.724)	(6.867)
Uncertainty(w)	21.01	6.010	21.67	5.456
	(27.32)	(17.90)	(18.67)	(16.60)
Interest rate	-0.109	0.0495	-0.0625	0.0349
	(0.113)	(0.0742)	(0.0774)	(0.0688)
Unemployment rate	0.0394	0.0338	0.00719	0.0465*
	(0.0421)	(0.0276)	(0.0288)	(0.0256)
Recession	0.447	0.449	0.373	0.140
	(0.491)	(0.322)	(0.335)	(0.298)
Constant	-0.603	-0.406	-0.385	-0.408
	(0.742)	(0.486)	(0.507)	(0.451)
Observations	100	100	100	100
R-squared	0.115	0.049	0.033	0.046

(6) The regression result of return of asset prices on sentiment scores of minutes(Minute_document level_Daily Data)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) SP_500 textual analysis	(2) eurusd textual analysis	(3) jpyusd textual analysis	(4) chfusd textual analysis
nt0	2.495	-2.365	-1.225	-1.379
	(2.348)	(1.587)	(1.607)	(1.456)
nt1	2.623**	-0.309	0.266	0.0703
	(1.268)	(0.858)	(0.868)	(0.786)
nt2	-0.900	0.0365	-0.579	-0.0822
	(1.089)	(0.736)	(0.745)	(0.675)
nt3	-2.989**	1.150	2.171**	0.487
	(1.287)	(0.870)	(0.881)	(0.798)
nt4	-1.380	-0.729	-0.298	-0.731
	(1.174)	(0.794)	(0.804)	(0.728)
nt5	2.200	-0.650	-3.073**	-0.470
	(1.791)	(1.211)	(1.225)	(1.110)
nt6	2.034	1.398	-0.321	0.700
	(1.297)	(0.877)	(0.888)	(0.804)
nt7	2.297	-1.210	-0.274	-0.962
	(1.870)	(1.264)	(1.280)	(1.160)
un0	0.460	3.688*	-1.383	3.680*
	(3.192)	(2.158)	(2.184)	(1.979)
un1	-1.998	-0.933	-0.496	-2.723
	(2.739)	(1.852)	(1.875)	(1.699)
un2	0.481	-0.162	1.372	1.257
	(2.136)	(1.444)	(1.462)	(1.325)
un3	4.211**	-0.990	0.820	-0.598
	(1.759)	(1.189)	(1.204)	(1.091)
un4	2.739	0.925	0.378	1.428
	(2.820)	(1.906)	(1.930)	(1.748)
un5	-3.125*	-0.544	-1.492	-1.486
	(1.641)	(1.109)	(1.123)	(1.017)
un6	-0.100	0.245	-0.473	-0.166
	(1.507)	(1.019)	(1.031)	(0.934)
un7	-3.060	-0.884	-2.965	-0.146
	(3.855)	(2.607)	(2.638)	(2.391)
Interest Rate	0.0621	0.139	-0.102	0.0686
	(0.167)	(0.113)	(0.114)	(0.104)
Unemployment	0.0780	-0.0191	0.0107	0.00557
	(0.0659)	(0.0446)	(0.0451)	(0.0409)
Recession	0.108	0.544	0.940**	0.158
	(0.532)	(0.360)	(0.364)	(0.330)
Constant	-1.627	-0.181	0.656	-0.174
	(1.054)	(0.712)	(0.721)	(0.653)
Observations	100	100	100	100
R-squared	0.303	0.202	0.236	0.216

(6) The regression result of return of asset prices on sentiment scores of minutes(Minute_Topic level_Daily Data)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	sp500 textual analysis	eurusd textual analysis	jpyusd textual analysis	eurusd textual analysis
nt0	1.044	-2.049	-1.061	-2.648
	(3.825)	(2.118)	(2.143)	(2.438)
nt1	3.469*	0.235	0.0144	-0.141
	(2.052)	(1.136)	(1.150)	(1.308)
nt2	-0.399	-0.706	-0.294	-0.879
	(1.646)	(0.912)	(0.922)	(1.050)
nt3	-4.392**	0.330	1.029	1.868
	(2.129)	(1.179)	(1.193)	(1.357)
nt4	-0.528	-0.598	0.00712	-1.008
	(1.717)	(0.951)	(0.962)	(1.094)
nt5	2.401	-1.467	-4.183***	-1.677
	(2.473)	(1.370)	(1.386)	(1.577)
nt6	1.684	0.479	-0.571	1.749
	(1.776)	(0.983)	(0.995)	(1.132)
nt7	0.306	-0.559	-1.104	-0.535
	(2.631)	(1.457)	(1.474)	(1.677)
un0	-1.884	4.412	-4.367	5.020
	(5.337)	(2.955)	(2.990)	(3.402)
un1	0.878	-2.277	-1.578	0.0497
	(4.034)	(2.234)	(2.260)	(2.572)
un2	-1.309	1.616	0.716	-0.535
	(3.417)	(1.892)	(1.914)	(2.178)
un3	5.752*	-1.366	2.590	-2.418
	(3.250)	(1.800)	(1.821)	(2.072)
un4	-1.221	-0.321	-1.399	0.790
	(5.361)	(2.969)	(3.004)	(3.418)
un5	-4.583	0.292	-2.660	1.658
	(2.838)	(1.572)	(1.590)	(1.809)
un6	1.181	-0.226	-0.625	0.576
	(2.368)	(1.311)	(1.326)	(1.509)
un7	-0.513	0.561	-0.802	-1.744
	(5.670)	(3.140)	(3.177)	(3.615)
Interest Rate	0.296	0.0498	0.0338	0.0535
	(0.284)	(0.157)	(0.159)	(0.181)
Unemployment	0.215*	0.0517	0.0780	-0.00465
	(0.119)	(0.0659)	(0.0667)	(0.0759)
Recession	0.329	0.225	1.132***	0.597
	(0.651)	(0.361)	(0.365)	(0.415)
Constant	-3.554*	-0.442	0.127	-0.0600
	(1.843)	(1.020)	(1.032)	(1.175)
Observations	68	68	68	68
R-squared	0.335	0.322	0.445	0.273

(7) The regression result of return of asset prices on sentiment scores of minutes(Minute_Topic level_ZLB_Daily Data)

VARIABLES	(1) sp500 textual analysis	(2) eurusd textual analysis	(3) jpyusd textual analysis	(4) eurusd textual analysis
Net tone(w)	14.06	-20.83**	-11.87	-14.43
	(15.12)	(8.098)	(9.717)	(9.484)
Uncertainty(w)	8.303	4.724	25.48	7.326
	(38.48)	(20.61)	(24.73)	(24.14)
Interest rate	-0.131	0.00733	-0.0208	0.00513
	(0.139)	(0.0746)	(0.0895)	(0.0874)
Unemployment rate	0.107	0.123***	0.0600	0.0817*
	(0.0684)	(0.0367)	(0.0440)	(0.0429)
Recession	0.370	0.205	0.492	0.481
	(0.569)	(0.305)	(0.366)	(0.357)
Constant	-0.766	-0.765	-0.834	-0.611
	(1.042)	(0.558)	(0.669)	(0.653)
Observations	68	68	68	68
R-squared	0.141	0.181	0.058	0.092

(8) The regression result of return of asset prices on sentiment scores of minutes(Minute_Document level_ZLB_Daily Data)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1