University of Warsaw
Faculty of Economic Sciences

Working Papers
No. 33/2020 (339)

THE IMPACT OF THE CONTENT OF FEDERAL OPEN MARKET COMMITTEE POST-MEETING STATEMENTS ON FINANCIAL MARKETS – TEXT MINING APPROACH

Ewelina Osowska
Piotr Wójcik

Warsaw 2020

UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
WORKING PAPERS

The impact of the content of Federal Open Market Committee post-meeting statements on financial markets – text mining approach

Ewelina Osowska[1], Piotr Wójcik[2] *

[1] *Data Science Lab WNE UW, University of Warsaw*
[2] *Faculty of Economic Sciences and Data Science Lab WNE UW, University of Warsaw*
* *Corresponding author: pwojcik@wne.uw.edu.pl*

Abstract: This article examines the impact of FOMC statements on stock and foreign exchange markets with the use of text mining and modelling methods including linear and non-linear algorithms. Proposed methodology is based on calculating the FOMC statements' tone called as sentiment and incorporate it as a potential predictor in the modelling process. Additionally, we incorporate the market surprise component as well as two financial indicators namely Purchasing Managers' Index and Consumer confidence index that gauge for corporate managers and retail customers assessment of the economic situation and potential fluctuations. Eight event windows around the event are considered: 60-minute and 20-minute windows before the event and also 15-minute, 20-minute, 25-minute, 30-minute, 60-minute and 120-minute windows after the event. Research has shown that given linear models the sentiment of FOMC statements does not generate a significant response in any of the analyzed event windows neither for the S&P 500 Index nor for the spot price on the EUR/USD currency pair. However, significant predictors occurred to be market shock in case of both S&P 500 Index and EUR/USD spot price, PMI in case of EUR/USD spot price and also CCI in case of EUR/USD spot price. Given non-linear models, the negative relation of statement's sentiment score and the model prediction is observed for EUR/USD spot price.

Keywords: FOMC Statements, event arbitrage, sentiment analysis, financial markets prediction

JEL codes: G14, G15, C14

Acknowledgements: Research presented in this article was partially financed by the Horizon 2020 FIN_TECH project (Grant agreement ID: 825215).

1. Introduction

Predicting financial markets behavior has always been an important topic to researchers. Traders can decide if to sell or buy a particular asset with the advance of time based on number of arbitrages among which still developing is event arbitrage. Event arbitrage is a high-frequency strategy which trades on the market movements surrounding an event. Events utilized in event arbitrage strategies might be any release of news about the condition of the economy, industry- or corporate-specific announcements, disturbances on the market, and other that impact market prices.

As an event of interest in this article, the release of Federal Open Market Committee (FOMC) public statements is considered. FOMC meets about eight times a year at the scheduled meetings to discuss monetary policy changes, decide on the federal funds target rate level, review economic conditions, assess price stability and employment situation. There are also conference calls and non-scheduled meeting conducted on non-regular basis[1]. For traders, FOMC meetings are time of increased volatility because any change in federal fund rates can affect a range of economic aspects. Also, besides the announced interest rate level, the content and specific wording is broadly analyzed in order to capture the tone of communication, FOMC's views on the economy and its policy propensity.

FOMC communication includes post-meeting statements (released immediately after the decision was made), speeches by the Chair, Chair's quarterly press conference, meeting minutes, summary of Economic Projections, Governors' speeches, congressional testimony by the Chair, speeches by bank presidents, Fed's semi-annual written report to Congress and news feed of the Fed[2]. There is a broad literature that investigates market reaction to FOMC communication among which the most commonly examined are meeting minutes. This paper also focuses on FOMC meeting minutes, however it broadens the existing literature on market reaction to statements.

Several studies have documented the fact that FOMC decisions have a significant impact on financial markets. Prior works on these issues focused on measuring the impact of

[1] These unscheduled meetings or conference calls happen always if there is a need for immediate action. For example, in 2020 there were two unscheduled meetings on March 3 (statement released at 10:00 a.m. EST) and March 15 (statement released at 5:00 p.m. EDT).

[2] All listed form of communications can be found on The Federal Reserve website: https://www.federalreserve.gov/newsevents.htm. Summary of Economic Projections (SEP) are released four times a year followed with a press conference by the chair. The meeting minutes of the scheduled meetings are released three weeks after the date of the policy decision (this state is actual from February 2, 2005).

Fed actions just through the fact of occurrence of meetings or releasing any announcements through the whole FOMC cycle (Cieslak, Morse and Vissing-Jorgensen, 2019), on the day of FOMC meeting (Bernanke and Kuttner, 2005; Kohn and Sack, 2004; Hayo, Kutan and Neuenkirch, 2008) or based on intraday stock reaction (Farka and Fleissig, 2013; Chirinoko and Curran, 2005).

Besides the fact of examining the impact of occurrence of any FOMC event on financial markets, other researchers also examined the content of FOMC communication in order to measure the tone and find its relationship to market reaction (Boukus and Rosenberg, 2006; Lucca and Trebbi, 2009; Cannon, 2015; Mazis and Tsekrekos, 2017; Jegadeesh and Wu, 2017).

In this paper we study in particular the effects and direction of influence of FOMC post-meeting Statements on financial markets returns. In order to measure this effect, we extract tone ("sentiment") of the statements and include it as an independent variable in estimated linear and non-linear models. We also account for market expectations regarding Federal Funds Target Rate using 30-days Federal Funds Futures. As additional predictors, we examine the Purchasing Managers' Index (PMI) and Consumer Confidence Index (CCI). Econometrically, we employ three modeling algorithms namely linear regression, support vector regression and random forest. The sample under investigation is from 2006 to 2020. In our view, this is a useful choice as it does include two global crises, namely financial crisis and coronavirus outbreak as well as many variations in the personal formation of FOMC.

We examine the stock and foreign exchange market reaction to the FOMC post-meeting statements release in order to validate three research hypotheses. First, information content of FOMC statements significantly influences financial markets reaction directly after the event. Second, additional control variables, i.e. surprise component, PMI and CB Consumer Confidence are significant predictors of the analyzed rates of return. Third, applying non-linear models results in better prediction of market reaction due to explaining more of data variability and taking into account potentially non-linear relationships.

The rest of this paper is structured as follows. Section 1 details the FOMC statements information content that is analyzed in context of market reaction. Section 2 details data issues, including data collation, sources, features preparation and its analysis. Section 3 presents the empirical strategy, including sentiment retrieval and modelling techniques in order to examine the relationship between market reactions to FOMC statement sentiment, market surprise component and other indicators. Section 4 provides the results of analyses and also introduces the ideas of further extending of the research. The last section summarizes and concludes the paper.

2. FOMC Statements and their impact on financial markets

In recent years, public statements issued by the FOMC have become an increasingly important component of US monetary policy. From the early 1980s, the FOMC holds eight regularly scheduled meetings per year, during which members discuss the economic outlook and formulate monetary policy. The FOMC first announced the result of a meeting including Fed Funds Target Rate decision on February 4, 1994[3]. Before 1994, the market instead deduced changes to the target of Federal Funds Rate from open market operations. In January 2000, the Committee announced that it would issue a statement following each regularly scheduled meeting, regardless of whether there had been a change in monetary policy[4]. Statements are published since then on the same day the policy decisions are made.

As to measuring FOMC communication content, similar text mining methodology namely Latent Semantic Analysis was applied by Mazis and Tsekrekos (2017) and Boukus and Rosenberg (2006). The former analyzed the FOMC statements and find that it incorporates many recurring topics. The themes appear in the study to be statistically significant in explaining the variation in 3-month, 2-year, 5-year and 10-year Treasury yields. In the study they control for monetary policy uncertainty by including information on expected Fed Funds Rate and concurrent economic outlook. The latter apply latent semantic analysis to FOMC minutes. Their analysis shows that the minutes reflect complex economic information and that they can be useful as a source of information in predicting economic activity. They discover that several themes are related to current market indicators, especially three-month yields and GDP growth.

Jegadeesh and Wu (2017) extend the analysis of extracting FOMC minutes topics by measuring each topic sentiment using lexicon-based approach. They then examined the informativeness of the overall document and with distinction to extracted topics meaning if market volatility increases following minutes release. They find that on the document level, market does not indicate significant reaction to neither positive sentiment nor negative sentiment. As to document topic level, they find that the market reacts significantly to monetary

[3] Federal Reserve press release on February 4, 1994 (https://www.federalreserve.gov/fomc/19940204default.htm).
[4] Poole, W., & Rasche, R. H. (2003). The impact of changes in FOMC disclosure practices on the transparency of monetary policy: are markets and the FOMC better "synched"? Federal Reserve Bank of St. Louis Review, 85 (https://files.stlouisfed.org/files/htdocs/publications/review/03/01/PooleRasche.pdf).

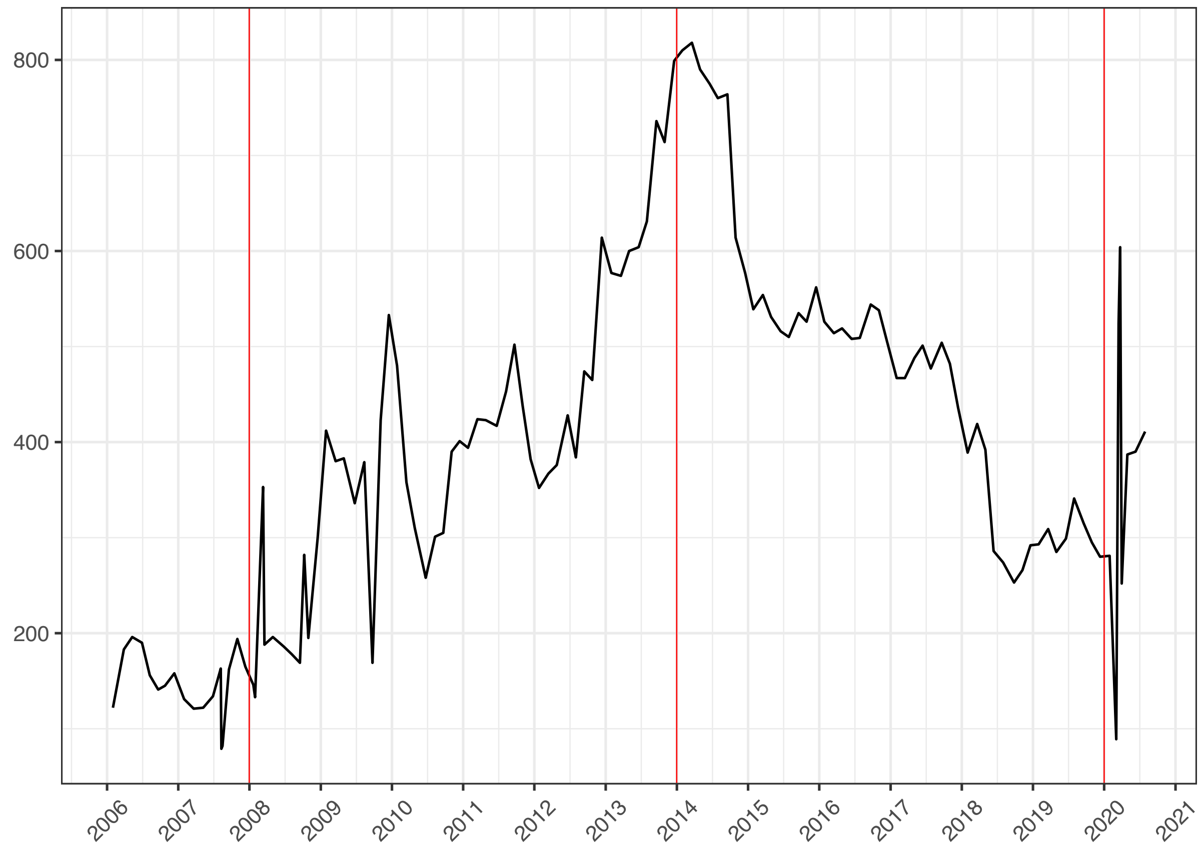
policy, inflation and employment and do not pay particular attention to growth-related discussions. As an equity market approximation, they use SPDR exchange-traded fund and S&P E-Mini futures contracts and as a bond market approximation they use Eurodollar futures contracts.

Cannon (2015) on the other hand focuses only on sentiment extraction and in order to score the content of FOMC meeting transcripts, they use lexicon-based approach for sentiment extraction using the combination of finance and consumer dictionaries. He discovers that sentiment of Committee discussions is strongly related to real economic activity approximated by the Chicago Fed National Activity Index (CFNAI) and also that the relationship varies by speaker class, more specifically the tone of Bank Presidents occurred to be more positive that the tone of the Governors and staff and also the tone of the staff occurred to be more positive that the tone of the Governors. His research regards the period between 1977 and 2009.

The content of communication was investigated also by Lucca and Trebbi (2009). They use two automated scoring algorithms regarding public statements: Google semantic orientation score and Factiva semantic orientation score. They find that for high-frequency data short-term Treasury yields respond to unexpected policy rate decisions and that longer-dated Treasury yields react mainly to changes in the content of communication.

Based on reviewed literature, we identified several research gaps which we decided to fill. Firstly, available literature provides mainly analyses based on basic methods which assume linear relationships between variables such as correlation or linear regression. In our research, both linear and non-linear methods are utilized in order to assess whether non-linear methods result in better prediction. What is more, another extension in our research regards taking into account additional control variables that are assumed to influence the financial markets. Lastly, as to exact statement release time, it is commonly assumed in many studies that statements are announced around 2:15 p.m. However, while it holds for the major part of statements, not all of them are announced at that time. Statements release time varies from meeting to meeting and what is more, the response of asset prices depends essentially on the time when market participants obtain the information. We follow the research by Rosa (2012) who obtained the FOMC statements announcement times by searching through several financial media sources to record the time the public first learned about the FOMC decision. The data regarding the period after 22.06.2011 (the last record in Rosa research) is sourced from investing.com platform.

Figure 1. Public statements length over time



Notes: The figure shows FOMC public statements length over time as measured by number of words in each individual statement. The sample runs from January 2006 to July 2020 and covers 123 statements. The part of statement regarding voting is not considered.

The general objective of this paper is to find whether there is a significant relationship between FOMC public statements' content and financial markets. We aim to verify three research hypotheses. First, information content of FOMC statements significantly influences financial markets reaction directly after the event. Second, additional control variables, i.e. surprise component, PMI and CB Consumer Confidence are significant predictors of the analyzed rates of return. Third, applying non-linear models results in better prediction of market reaction due to explaining more of data variability and taking into account potentially non-linear relationships. We decided to conduct research considering two financial markets namely stock market and foreign exchange market due to the fact that several studies (such as Hayo, Kutan and Neuenkirch, 2008 or Cieslak, Morse and Vissing-Jorgensen, 2014) proven these markets indicate significant reaction to FOMC communication. On the stock market, we analyze returns of S&P 500 global index which includes 500 US companies with the highest capitalization and thus is considered to be the best single equity that gauges the biggest companies' performance.

As to foreign exchange market, we analyze EUR/USD which is the most important currency pair in the world that significantly reacts to both Europe and US events.

In recent years, the FOMC has improved its communication with the public. Nowadays, the Fed reports more frequently and more comprehensively on the economic matters. The public statements form evolved over years regarding the length, provided information and wording. The period analyzed in this article covers only the newest regime (years from 2006 to 2020) and thus only changes in statements applying to that period will be described. As to the statements' length, *Figure 1* indicates the number of words used in the statements through the whole analyzed period of time. Significant changes occurred after 2008 (financial crisis), 2014 (Janet Yellen takes over the role of Chairman) and 2020 (coronavirus outbreak).

Besides the statement's length, also the content has changed. At the beginning of the analyzed period, statements included only general information on decided Federal Funds Target Rate, current assessment of economic situation, likely future course of economic situation and the list of participants voting for a specific action. With time, statements became more informative and transparent. The current form of the statement includes six main features. It begins with the assessment of the state of the current U.S. economy since the last FOMC meeting. After that the monetary policy goals are emphasized. The following statement regarding FOMC objectives is recurring in almost every statement: *"Consistent with its statutory mandate, the Committee seeks to foster maximum employment and price stability"*. Later in the statement, decision regarding Federal Funds Target Rate is presented. After the policy decision, likely future course of economic situation is introduced. The announcement incorporates additionally what factors the FOMC plans to consider in coming strategy choices. For an example FOMC section see *Appendix A*. The last element of the statement is the list of participants voting for the FOMC monetary policy action. It is important to note that the statements are rather uniform and contain identical sentences with the only difference being the assessment of economic situation. Several statements from the analyzed period that reflect described changes are included in the *Appendix B*.

3. Methodology

3.1. Modelling approach

We examine the relationship between the information content of FOMC statements, market projections, Purchasing Managers' Index, Consumer Confidence Index and returns of S&P 500 Index and EUR/USD to assess the influence of FOMC actions on financial markets. In order to obtain the relation, we estimate three models using such algorithms as linear regression, support vector regression and random forest. These methods were chosen mainly due to its popularity and proven good performance. Linear regression model is a baseline method in case of many researches and what is more, for almost each research regarding FOMC communication impact on financial markets. As to support vector regression and random forest these are two non-linear methods from the area of machine learning, based on very different assumptions. SVR aims to find a hyperplane fitting the data after applying some non-linear transformations based on a selected kernel function while random forest is one of the most popular ensemble tree-based method.

Several studies examined the market reaction within event windows that occurred to be significant to other researchers. For example, Schumaker et al. (2012) selected 20-minute event window following Gidofalvi (2001) who observed 20-minute window of weak predictability. As to researches strictly related to FOMC communication, Boukus and Rosenberg (2006) considered only 20-minute event window and Farka and Fleissig (2013) considered 30-minute interval. In our analysis, we compare the market reaction to chosen indicators with respect to eight specified event windows around the time when the post-meeting statements are released. We consider two pre-event windows, namely an hour and half an hour before statement is released and five post-event windows, namely 15-minute, 20-minute, 25-minute, 30-minute, an hour and half an hour after statement is released. Pre-event windows are considered in order to examine if there is any insider effect visible on the markets. It happens if buying or selling decisions of insiders are followed by other investors and thus market reacts in advance of an actual event.

Based on specified event windows, we then calculate returns as absolute value of the difference between an asset value at time of FOMC statement release and an asset value at the end of event window. The following formula presents the S&P500 Index returns calculation in case of 15-minute event window:

$$R_{\cdot 2:15}^{SP500} = \frac{Y_{2:15}^{SP500} - Y_{2:00}^{SP500}}{Y_{2:00}^{SP500}}, \tag{1}$$

where $Y_{2:15}^{SP500}$ is the close price of S&P500 Index reported at 2:15 p.m. and $Y_{2:00}^{SP500}$ is the close price of S&P500 Index reported at 2:00 p.m.

We follow Kohn and Sack (2004), Bernanke and Kuttner (2005) and Farka and Fleissig (2013) and gauge for market policy expectations. It is clear that expectations of Fed policy actions are not directly observable. However, according to Kuttner (2001), Fed funds futures prices are proper market-based approximation for those expectations. We thus utilize the approach proposed by Kuttner (2001) and adopted also by earlier mentioned researchers. For each monetary policy announcement, we calculate the unexpected component of FOMC decisions as follows:

$$\Delta r_t^u = \frac{m}{m-d}(f_t - f_{t-1}) \tag{2}$$

where m is the number of days in month, d is the day of FOMC statement release and f_t (f_{t-1}) are the futures rates at time t ($t-1$). If there is no statement release on a particular day or the statement is released on the last day of month, the surprise component of zero is assigned.

The general model specification is as follows:

$$R_t = Score_t + \Delta r_t^u + PMI_t + CCI_t + \varepsilon_t, \tag{3}$$

where $Score_t$ is the tone of statement released at time t, Δr_t^u is an unexpected component of FOMC decisions calculated based on 30-day Federal Funds Futures, PMI_t is value of Purchasing Managers' Index holding at time t, CCI_t is a value of Consumer Confidence Index holding at time t and $\in_t$ is prediction error component. The rationale behind considering both PMI and CCI indexes is that it controls for the attitudes toward economic situation from two perspectives namely business and individual customers.

For all of these event windows we estimate a linear regression which we consider as a baseline model, support vector regression and random forest. Linear least squares regression is by far the most widely used modeling method in many different research areas. It is a simple and universal statistical method used for finding linear dependencies between continuous dependent variable and several features called independent variables. The general objective of this approach is to minimize the error rate, more specifically to minimize the sum of the squares of the residuals. This method enables to estimate unknown coefficients which inform to what extend and in which direction independent variables influence the dependent variable. Based on the estimated coefficients and the actual values of independent variables, one is able to obtain forecast of the dependent variable.

Support vectors are one of the most popular tools used for both classification (SVM) and regression tasks (SVR). In a linear problem, SVM performs classification by finding a hyperplane that maximizes the margin between observations that belong to different classes. In case of non-linear problem, the kernel function transforms the data into a higher dimensional feature space to make it possible to perform the linear separation. There are several kernel functions that can be used for that problem. The most popular ones are a linear kernel, a polynomial kernel and a Gaussian radial basis function. The concept of SVR is very similar to SVM, however it requires additional assumptions. SVR aims to fit the best regression hyperplane within a certain threshold of error value ε. This approach aims to minimize the effect of outliers. One of the main advantages of SVR is that its computational complexity does not depend on the dimensionality of the input space.

The last algorithm, random forest, is an ensemble learning method that builds a number of different weak prediction models which in our case are regression trees. The final model prediction consists of the results of all weak learners. The so-called weak learner is a model which is slightly better than random guessing. The way the final result is obtained depends on the type of ensemble method. Most common ensemble types are bootstrap aggregating (bagging), boosting and stacking. The main advantage of using ensemble methods over simple decision trees (classification task) or regression trees (regression task) is that it results in better predictive performance than could be obtained from any of the weak learners alone and also corrects for trees' tendency of overfitting to their training data set. Random forest utilizes bagging, which repeatedly selects a random sample with replacement (bootstrap) from the training data set and independently fits trees to these samples. The final result is the average of all model predictions made on all selected subsamples. An important characteristic of random forest is that, at each split in each individual tree it uses a random subset of predictors. This prevents individual trees from being correlated because if one or a few features appeared to be strong predictors, they would be selected in many trees in the top split resulting in very similar trees. Averaging results of similar trees would not reduce variance as much as averaging results of uncorrelated trees. Thus, this limited number of considered features reduces that risk. Random forest requires hyperparameters optimization including in particular a number of trees to build in the forest and number of variables randomly sampled at each split.

It is important to note that we do not consider in the modelling statements released on March 28, 2006 (2:17 PM), August 17, 2007 (8:15 AM), March 11, 2008 (8:30 AM), October 8, 2008 (7:00 AM) and March 23, 2020 (8:00 AM). These are mostly the statements released on early morning hours for which there is no data on assets quotations available from our source.

For each model, the Root Mean Squared Error (RMSE) measure was used as a loss function while model tuning. In case of non-linear models such as SVR and random forest, results were obtained using cross-validation procedure which served the purpose of additional model validation. We use 10-fold cross validation and obtain 100 model estimates. The final result is an average of each of these 100 models. This approach enables for more accurate prediction, especially in case of low-numerous samples as in the following research. What is more, we additionally separate the validation sample consisting of 12 most recent observations in each modelling sample for the purpose of non-linear models estimation. We use obtained validation samples to calculate additional performance statistics on the data the model hasn't seen while training. Obtained models are validated with the use of such statistical measures as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The significance of features in linear regression model is assessed using the p-value. In case of non-parametric machine learning models, the variables importance is obtained with the use of Explainable Artificial Intelligence (XAI) method, namely Permutated Feature Importance (PFI). PFI is calculated based on the training sample. Additionally, in order to assess the relationship between the prediction and independent features in non-linear models, the analysis of Partial Dependence Plots (PDP) is conducted.

3.2.Initial text preprocessing

As an initial text processing steps before statement's sentiment quantification, the following text cleaning actions were undertaken. First, all downloaded statements were converted into the Corpus which is a standard text representation in Natural Language Processing. After that all words were lowercased. Additionally, punctuation marks and numbers were removed. Later words from list of English stop words were excluded. The list was originally sourced from tm package implemented in R Software. The list was extended with FOMC abbreviation expansion and also stop words list provided by Loughran and McDonald[5]. Then all words were subject to the process of lemmatization. The goal of lemmatization is to reduce inflectional form of a word to a base form. Finally, the tokenization took place which is a process of separating a piece of text into smaller units. In case of our analysis these units are words. The fragment of FOMC

[5] Available at https://sraf.nd.edu/textual-analysis/resources.

statement released on January 31, 2006 before and after initial text processing is presented in Appendix C.

3.3.Sentiment extraction approach

There are two primary methods of calculating text sentiment including unsupervised and supervised methods. The major difference between these two types of algorithms is that unsupervised algorithms base on non-tagged data while supervised algorithms based on labeled data. In case of FOMC Statements, no tags are officially available and thus the problem comes down to unsupervised learning. The benefits of using unsupervised learning is that no human-tagged training data is needed which is costly and subjective, however, it usually doesn't perform as well as supervised learning.

There were several attempts to label FOMC communication including human tagging (Hayo, Kutan and Neuenkirch, 2008), market reaction to release (Zadeh and Zollmann, 2009) and statistical methods (e.g. Boukus and Rosenberg, 2006; Cannon, 2015; Jegadeesh and Wu, 2017). There are several statistical approaches used to extract the content of FOMC statements. For example, Boukus and Rosenberg (2006) and Mazis and Tsekrekos (2017) utilize Latent Semantic Analysis in order to extract distinct topics which are covered in post-meeting statements. Topics were extracted also by Jegadeesh and Wu (2017) with the use of algorithm called Latent Dirichlet Allocation. Besides extracting topics, there were also attempts to extract the tone of the communication by using sentiment analysis. With the use of lexicon-based sentiment analysis approach, Jegadeesh and Wu (2017) calculated sentiment for each extracted topic and Cannon (2015) calculated sentiment of the overall document. Lucca and Trebbi (2009) on the other hand, use automatic algorithms implemented by Google and Dow Jones (Factiva tool).

In this study, we follow Cannon (2015) and Jegadeesh and Wu (2017) and utilize the lexicon-based approach for sentiment calculation. In this approach, the sentiment of a statement depends on the sentiment of each individual term that composes it. There are several dictionaries available for this task including general and domain-specific dictionaries. We decided to create a hybrid solution which consists of the Loughran-McDonald financial dictionary[6] and Harvard IV-4 psychological dictionary[7]. We merge these two dictionaries in order to extend the list of words with the assigned sentiment. Alternative domain-specific

[6] Available at https://sraf.nd.edu/textual-analysis/resources.
[7] Available at http://www.wjh.harvard.edu/~inquirer/homecat.htm.

dictionary is Henry's finance dictionary proposed in 2008. However, it includes a very limited list of positive and negative words in comparison to Loughran-McDonald dictionary and thus we decided not to consider it. It is important to note that we treat the Loughran-McDonald dictionary as a dominant one since it also includes words from Harvard IV-4 dictionary but accounts for financial context and thus the sentiment differs in some cases. We also modify the tone value assigned to word "inflation" included in Harvard IV-4 dictionary since it is recognized as negative. Instead we assign a neutral value due to the fact that the word itself does not carry a negative meaning in the context of FOMC statements. Based on the described customized dictionary, once a text is initially prepared as described earlier in this section, we calculate the sentiment score for each statement using the following formula:

$$Score_t = \frac{positive_t - negative_t}{positive_t + negative_t}\left(\frac{1}{N_t}\right), \tag{4}$$

where $positive_t$ is the number of positive words in the statement released at time t, $negative_t$ is the number of negative words in the statement released at time t and N_t is the total number of words in the statement released at time t. Score above zero indicates positive tone, score below zero indicates negative tone while score equal to zero indicates neutral tone. We consider positive tone as rather optimistic outlook of the economy, while negative tone is related to poorer economic conditions.

There are several limitations of using lexicon-based approach for sentiment extraction. Firstly, as indicated before, this approach is usually outperformed by supervised algorithms. Secondly, it does not account for the context of the text. For example, words indicating an increase are in general considered to be positive. However, in the context of inflation or unemployment, increase indicates a negative phenomenon. It is not reflected in the score obtained with the lexicon-based approach. One solution to that problem is manual investigation and modification of the assigned sentiment with regard to the context or excluding words related to the increases and decreases.

Figure 4 presents the sentiment score of each statement released between January 31, 2006 and July 29, 2020. What can be seen immediately is that the sentiment is the most volatile at the beginning of the analyzed period which might be related to the fact that at that time FOMC statements were the shortest over the years. Also, sentiment trend clearly shows the financial crisis around 2008 and to a much lesser extent the coronavirus outbreak in 2020.

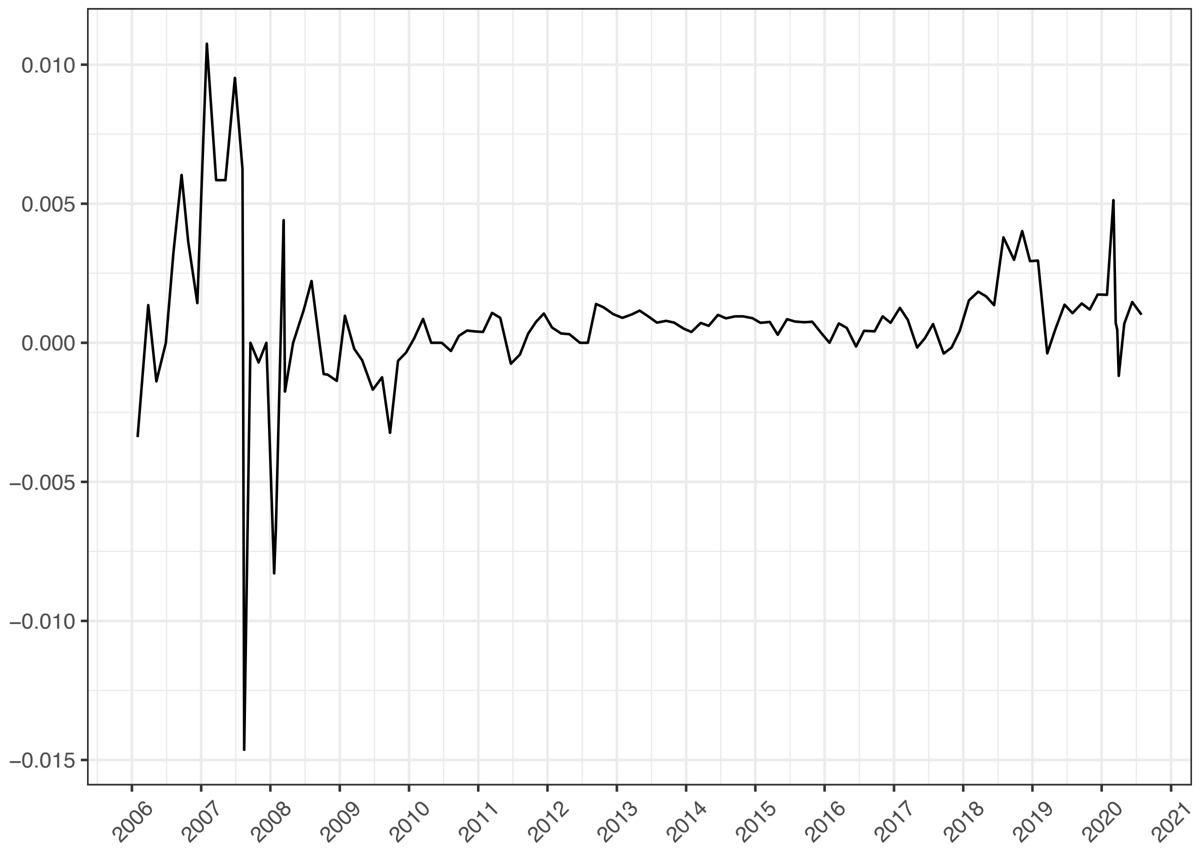


Figure 4. The FOMC Statements sentiment score obtained with lexicon-based approach

Notes: The figure shows the trend of net sentiment. Analysis includes 122 FOMC statements released between January 31, 2006 and July 29, 2020. Meeting held on January 21, 2008 was omitted due to lack of key financial data.

4. Data

4.1.FOMC Statements

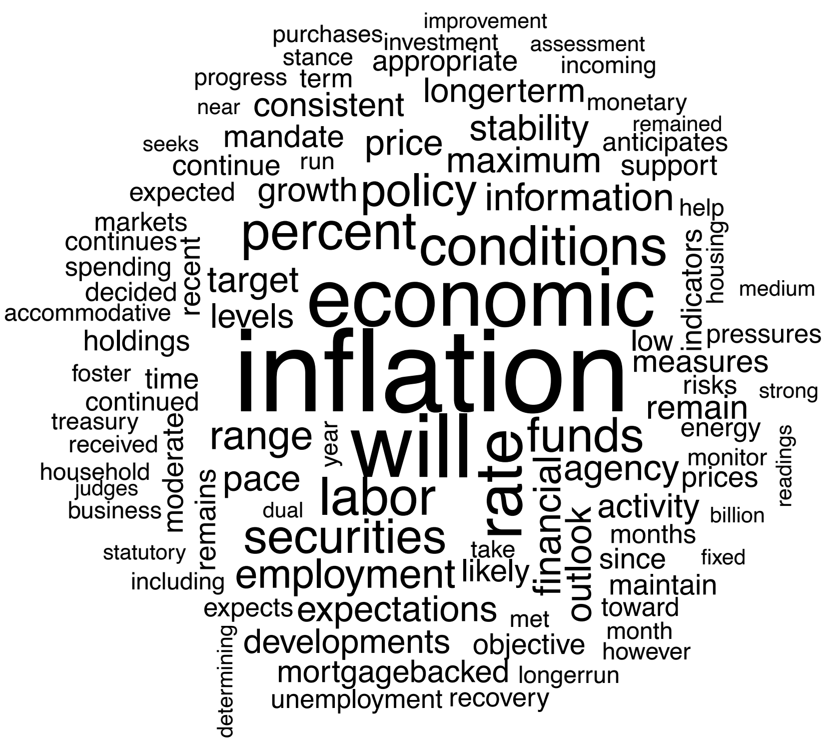
Our analysis is based on FOMC statements released over the period from January 2006 to July 2020. Total sample comprises of 123 statements. Statements were collected from Federal Reserve website[8] with the use of web-scraping algorithm implemented in Python language. In the scraping algorithm we limited the content only to FOMC announcement, omitting the names of voters listed in the last paragraph. For the purpose of the primary analysis, we consider

[8] FOMC Statements are available at https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm.

statements released after both scheduled and unscheduled meetings including conference calls. We additionally exclude two statements which were released on the weekend[9].

FOMC Statements are initially analyzed in terms of their length and most common words. Results of the analysis of statements length is presented earlier in the article. As to most common words, *Figure 2* presents a word cloud summarizing 100 most frequently terms used by the Fed. It can be clearly noticed that a term with the highest frequency is 'inflation'. Besides, common words are also 'economic', 'will',' rate', 'percent' or 'conditions'. Despite the fact that the figure was created after stop words removal, there are still present terms that do not carry any significant content such as all of those listed above. These terms will be naturally filtered out in the process of sentiment calculation since these are not available in utilized dictionaries.

Figure 2. Word cloud of most common words in FOMC statements in the period 2006-2020



Notes: The figure shows a hundred most commonly used words in FOMC public statements released between January 2006 and July 2020. Words were limited to those which occurred at least 20 times. Sample consists of 123 statements.

[9] Excluded statements were released on May 9, 2010 (Sunday) and March 15, 2020 (Sunday).

The data cleaning process reduced the number of words approximately by half. The average number of words in the statement before cleaning was 392, while the average number of words in the statement after cleaning was 195.

4.2.Financial market data

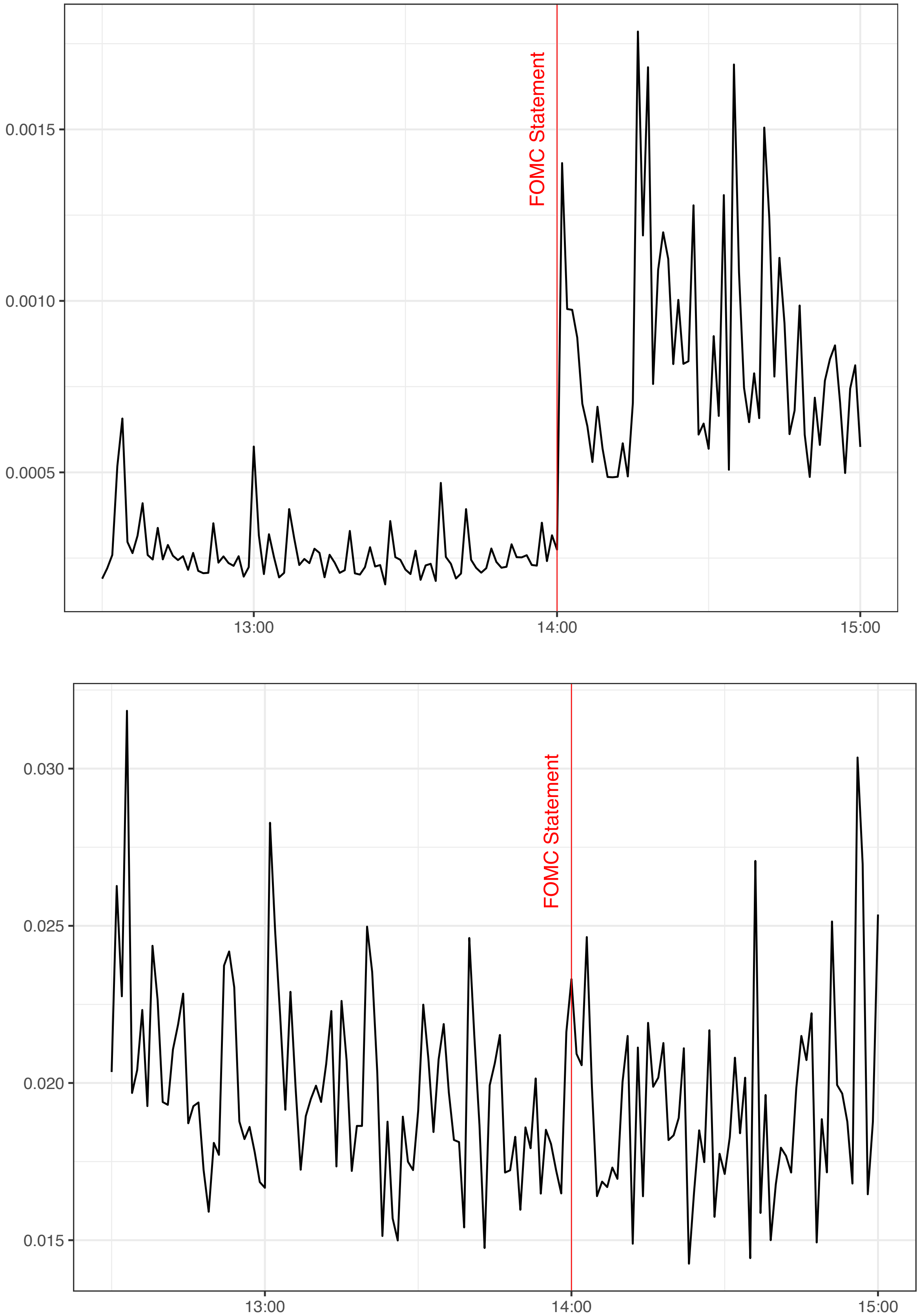
In order to gauge market reaction to FOMC statements, we use high-frequency quotations from both stock and foreign exchange markets. For the stock market, our main indicator is 5-minute intraday data on S&P 500 Index. We use the closing price of an index documented every 5 minutes. For the exchange rate market, we use data on EUR/USD spot rate. Data is collected from www.finam.ru website for the period from January 1, 2006 to July 31, 2020. *Figure 3* shows the reaction in terms of volatility of asset prices measured with standard deviation around FOMC statements releases. It is obtained as an aggregated standard deviation per one-minute time interval from all analyzed FOMC meeting days. The formula is as follows:

$$Volatilty_t = \sqrt{\frac{\sum_{t=1}^{T}(X_t - \bar{X})}{T-1}}, \tag{5}$$

where T is the number of observations in the sample, X_t is the one-minute return, $\bar{X}$ is the sample mean. Number of observations is usually equal to the number of FOMC meetings in scope of analysis, unless for some date there is no observation available at a specific point of time.

The vertical line indicates the approximate release time of the FOMC statement namely 2 p.m. As to S&P 500, a significant increase in index volatility is observed at a release time and it also persists after the release. Maximum volatility before release is about 0.00075 while after the release, maximum volatility reaches above 0.00175. In case of EUR/USD volatility, it is in general higher than S&P 500 volatility for the whole event window. No significant increase around FOMC statement approximate release time is observed. Instead, several volatility peaks can be seen during the observed window. The empirical analysis takes into account 123 FOMC meeting days between January 31, 2006 and July 29, 2020. One meeting held on January 21, 2008 was omitted due to the lack of S&P 500 data on that day, most probably resulted from the fact that that day was Black Monday in worldwide stock markets.

Figure 3. Volatility of S&P 500 (top panel) and EUR/USD (bottom panel) around statement release



Notes: The figure shows the standard deviation of S&P 500 *and EUR/USD* returns around FOMC statement release. The vertical red line indicates the approximated release time of the FOMC statements, that is, 2 p.m. EST. Analysis includes 122 FOMC meetings from January 31, 2006 to July 29, 2020. Analyzed window of time is from 12:30 p.m. until 3 p.m.

In order to correctly analyze the impact of FOMC statements on both assets, we first needed to map the event window of the announcement with respective financial data according to *Content of FOMC Statements*. In this section, we analyze eight different event windows that is 60-minute and 30-minute before the event and 15-minute, 20-minute, 25-minute, 30-minute, 60-minute and 120-minute after the event. For example, assuming the statement was announced at 2 p.m., 30-minute event window is obtained from 2 p.m. to 2:30 p.m., 25-minute event window is obtained from 2 p.m. to 2:25 p.m., 20-minute event window is obtained from 2 p.m. to 2:20 p.m. and 15-minute event window is obtained from 2 p.m. to 2:15 p.m. Table 1 summarizes statistics after mapping for high-frequency returns around policy announcements regarding four abovementioned event windows from January 1, 2006 to July 31, 2020.

Table 1. Summary statistics of returns of S&P 500 and EUR/USD on different intervals directly before and after FOMC statements release

Statistics	60-min before event	30-min before event	15-min after event	20-min after event	25-min after event	30-min after event	60-min after event	120-min after event
S&P 500								
Mean	-0.0002	0.0000	0.0004	0.0005	0.0004	0.0009	0.0013	0.0006
Std	0.0037	0.0024	0.0039	0.0043	0.0047	0.0053	0.0053	0.0058
Min	-0.0220	-0.0155	-0.0091	-0.0114	-0.0159	-0.0261	-0.0116	-0.0123
Max	0.0116	0.0081	0.0129	0.0161	0.0173	0.0226	0.0203	0.0176
EUR/USD								
Mean	0.0001	-0.001	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001
Std	0.0011	0.008	0.0005	0.0005	0.0005	0.0006	0.0008	0.0017
Min	-0.0051	-0.0032	-0.0022	-0.0019	-0.0017	-0.0016	-0.0022	-0.0039
Max	0.0029	0.0023	0.0019	0.0018	0.0013	0.0017	0.0017	0.0064

Notes: Statistics calculated for S&P500 and EUR/USD, based on 60-minute and 30-minute event windows before event and 15-minute, 20-minute, 25-minute, 30-minute, 60-minute and 120-minute event windows after event, including only FOMC statements release days.

4.3.Other indicators

Besides the FOMC statement quantifier and interest rate surprises, we additionally consider in the analysis two economic indicators namely Purchasing Managers' Index (PMI) and Consumer

Confidence Index (CCI). We sourced the actual monthly data as well as forecasts of these measures from the investing.com platform[10]. PMI is calculated based on a monthly survey of senior executives from about 400 companies in 19 industries. Survey is based on five areas: new orders, inventory levels, production, supplier deliveries, and employment. Questions are related to business conditions and any changes in the business environment regarding purchases. As to CB Consumer Confidence indicates consumer confidence level with respect to economic stability. The index is based on a monthly survey of about 3,000 households that are located all around the United States. The questionnaire includes five questions from two areas being current economic conditions and expectations regarding the changes in economic conditions.

5. Results

In this section, we examine the stock and foreign exchange market reaction to the FOMC post-meeting statements release in order to explore three research hypotheses. First, information content of FOMC statements influences financial markets. It comes down to a question if the content of FOMC statements matter to investors. Second, additional features namely surprise component, PMI and CCI are significant predictors of the analyzed rates of return. Third, estimating non-linear models results in a better prediction of market reaction due to explaining more of data variability. We fit linear regression, support vector regression and random forest for eight event windows calculated based on the time the statements were released, only on release days. Using quantified tone of the FOMC statements and values of control variables in the model estimation will allow us to verify the two first research hypotheses. The last hypothesis is verified by comparing the performance between the estimated linear and non-linear models.

We begin with the results of estimation of S&P 500 returns. Table 2 presents model coefficients estimated based on a linear regression. It is important to note that the modelling sample consists of about a hundred observations in case of event windows 60-minute and 30-minute before the event and also 15-minute, 20-minute, 25-minute, 30-minute and 60-minute after the event. For 120-minute event window, the number of observations is the lowest and totals to 53. Due to low numerosity of samples, the results should be interpreted and relied with caution. As to variable indicating the statement's sentiment score (*Score*), it did not occur to be

[10] Data on PMI are available at https://www.investing.com/economic-calendar/ism-manufacturing-pmi-173 and data on CB Consumer Confidence are available at https://www.investing.com/economic-calendar/cb-consumer-confidence-48.

significant in any linear regression model specification and thus it is difficult to explicitly indicate the impact of the tone and in particular the direction of this impact. Significant relationship can be seen however in case of a variable indicating the market surprise component (*Market shock*) for three event windows: 15-minute, 30-minute and 60-minute. For each event window the direction of shock influence is additive which means that with a higher market surprise component measured with Federal Funds Futures, the higher return is observed. As to variables PMI and CCI, similarly as in the case of variable indicating sentiment's score, no significant relationship is observed for all eight model specifications. Referring to the Figure 3, on the basis of a linear regression, we can associate a significant increase in the volatility of S&P 500 Index mainly with the effect of market surprise. What is more, based on the linear regression results, we are not able to state an insider effect on the market. For the record, the insider effect would be observed if significant relation of potential prediction and target variable was seen for windows 60-minute and 30-minute before the event.

Tables 3, 4, 6 and 7 report the results of the support vector regression and the random forest. Instead of model coefficients, as in the case of linear regression, the table contains the permutated feature importance of variables using RMSE as loss function. Permutation-based feature importance (PFI) is a measure of the impact of a feature on the target variable which provides for the main feature effect and also effects of interactions with other features. The value itself is interpreted as the increase in model error if a particular feature would not be included in the model.

Moving to direct interpretation of the results in Table 3, it occurs that in case of the SVR model, all variables represent some portion of importance regarding S&P 500 returns. For each model specification, there is no dominant variable that has significantly larger influence on the target variable than other features. What is more, permutated feature importance for every variable is very similar. The biggest difference of variables importance can be seen for 20-minute, 60-minute and 120-minute event windows. Random forest results are presented in Table 4. The conclusions are very similar as in case of SVR. However, there is a visible increase of the importance for a variable indicating market shock in case of 15-minute, 20-minute and 60-minute event windows. For 15-minute and 20-minute event windows, the increase in importance ties in with the significance of market shock variable indicated in the linear regression model. For non-linear methods in order to extract the direction of the relationship between target variable and predictors, Partial Dependency Plots (PDP) are obtained. Due to the fact that in our research we are focused mainly on the variable indicating statements' sentiment score, only this variable is investigated in the analysis of PDPs. The resulting figures can be found in *Appendix D.* Tables D1 regards S&P 500 Index returns models estimation and lists plots for each event window for both SVR and RF models. Based on these plots we will try to evaluate what is the direction of the influence of FOMC statements' sentiment score on the S&P 500 returns. We can observe several shapes of the dependence curves which indicate that different linear and non-linear relations are in fact observed between the score variable and the prediction outcome of the target variable. Linear relation results from using a linear kernel in SVR. It is not always clear what is the direction of dependence due to the fact that curves sometimes break in several places. Let us start with the SVR. In case of 60-minute before event window we observe the only case of additive relation of sentiment' score and the model prediction of S&P 500 returns. In other cases (30-minute before, 15-minute, 20-minute, 25-minute, 30-minute and 60-minute event windows), we observe mostly negative relation. For random forest in case of 60-minute window before the event the observed relation is additive. For other event windows, the relation is rather difficult to assess.

Tables 5-7 report the results of estimation of EUR/USD spot price. Table 5 presents model coefficients estimates based on the linear regression. Similarly as in case of S&P 500 Index, variable indicating the statement's sentiment score (*Score*) is not significant in any model specification. Significant relationship can be seen however in case of a variable indicating the market surprise component (*Market shock*) for three event windows: 60-minute before the event, 15-minute and 20-minute after the event. For three event windows, namely 15-minute and 20-minute after the event the direction of shock influence is negative which is counterintuitive. Only for 60-minute before event window the observed relation is additive which means that a higher market surprise component results in observing higher return. As to variables PMI and CCI, PMI is significant for 30-minute before and 120-minute after event windows and CCI is significant for 30-minute after event window. In case of EUR/USD spot price, we also cannot observe the insider effect, same as with S&P 500 Index. As to non-linear algorithms' results, in case of SVR (Table 6) all variables represent some portion of importance. Analogous to S&P 500 Index, there is no dominant variable and also the PFI values are very similar between variables and event windows. The only exception from that is the 120-minute event window in case of which the importance is about twice as high as in case of other event windows. The biggest difference of variables importance can be seen for 60-minute before and 15-minute after event windows for variable indicating market shock and for 30-minute after event windows for variable indicating PMI. Random forest results are presented in Table 7. The conclusions are very similar as in case of SVR. A significant difference in PFI can also be seen for 120-minute event window. It is also twice as high as in case of SVR. There is a visible increase of the importance for a variable indicating score for 60-minute before event window and for a variable indicating CCI for 60-minute and 30-minute before event windows.

Table D2 in *Appendix D* lists plots for each event window for both SVR and RF models regarding EUR/USD spot price models estimation. Similarly as in case of the S&P 500 returns, for EUR/USD we can observe many different shapes of dependence curves which results from using different modelling methods. It is not always clear what is the direction of dependence due to the fact that curves sometimes break in several places. However, the most common direction is negative. This means that most commonly (for the most event windows) the decline in the average predicted value of EUR/USD spot price is related to the increase of sentiment' score. More intuitively, the lower score causes larger reaction of the market, while a higher score causes smaller reaction of the market.

In order to verify whether non-linear models result in a better prediction than linear models, we compare the accuracy of predictions based on testing sample between models and also conduct the stability analysis of MAE and RMSE measures between the training sample and test sample. We observe generally lower difference in MAE and RMSE between training and test sample for linear regression compared to support vector regression and random forest for all considering event windows. However, regarding the forecast error, we observe on average lower values of out-of-sample MAE and out-of-sample RMSE for the support vector regression model. Based on that, it is difficult to verify the hypothesis unequivocally, because on the one hand the support vector regression model is slightly better than the others considered, but on the other hand, when analyzing the stability of the forecast quality, we observe that linear regression is doing better. Thus, we are not able to neither confirm nor reject the hypothesis that the non-linear models explain more of data.

Summarizing the above, the research allowed for verification of all research hypotheses. First, we cannot confirm the hypothesis that the information content of FOMC statements significantly influences financial markets reaction. Based on linear regression results, the variable indicating the tone of the statement did not occur to be significant for both S&P 500 Index and EUR/USD spot price. Based on non-linear models it was also not possible to confirm the hypothesis due to the fact that we cannot observe significant increase in importance of variable indicating statements' tone. Second, we are able to confirm the hypothesis that additional control variables namely surprise component, PMI and CB Consumer Confidence are significant predictors of analyzed rates of return. We confirm the hypothesis based on the obtained significant relationship in the linear regression regarding both S&P 500 Index and EUR/USD spot price. Third, we are not able to neither confirm nor reject the hypothesis that non-linear models result in better prediction of market reaction due to explaining more of data variability. We based our assessment on the comparison of forecast errors for linear and non-linear models and the stability analysis of MAE and RMSE measures between the training sample and test sample.

Table 2. Linear regression model results for S&P 500

Predictors	*60-minute before event*	*30-minute before event*	*15-minute after event*	*20-minute after event*	*25-minute after event*	*30-minute after event*	*60-minute after event*	*120-minute after event*
(Intercept)	0.004535	0.002273	-0.001937	0.003988	-0.010071	0.003898	-0.005507	0.005208
	(0.290)	(0.435)	(0.644)	(0.473)	(0.080)	(0.497)	(0.359)	(0.756)
Score	0.138626	-0.042225	-0.012030	-0.123136	-0.052377	0.056610	0.079734	-0.485619
	(0.522)	(0.773)	(0.957)	(0.613)	(0.836)	(0.847)	(0.770)	(0.802)
Score PFI	0.003707	0.002513	0.003557	0.004207	0.004444	0.005193	0.004821	0.005901
Market shock	0.000571	-0.000152	**0.003824**	0.002214	0.002127	**0.003106**	**0.004501**	0.003903
	(0.558)	(0.818)	(0.000)	(0.062)	(0.073)	(0.030)	(0.001)	(0.854)
Market shock PFI	0.003693	0.002501	0.004223	0.004427	0.004619	0.005512	0.005567	0.005872
PMI	-0.000064	-0.000040	0.000044	-0.000054	0.000184	-0.000057	0.000127	-0.000046
	(0.473)	(0.507)	(0.611)	(0.628)	(0.115)	(0.637)	(0.296)	(0.891)
PMI PFI	0.003717	0.002517	0.003567	0.004216	0.004611	0.005219	0.004866	0.005879
CCI	-0.000020	-0.000002	-0.000002	-0.000007	0.000005	-0.000004	-0.000003	-0.000015
	(0.250)	(0.852)	(0.882)	(0.720)	(0.802)	(0.881)	(0.907)	(0.764)
CCI PFI	0.003755	0.002507	0.003560	0.004220	0.004452	0.005192	0.004809	0.005885
Observations (training sample)	96	98	95	95	94	99	94	41
MAE	0.0021	0.0014	0.0025	0.0029	0.0030	0.0033	0.0034	0.0041
RMSE	0.0037	0.0025	0.0036	0.0042	0.0044	0.0052	0.0048	0.0059
Observations (out-of-sample)	12	12	12	12	12	12	12	12
OOS MAE	0.0023	0.0012	0.003	0.0031	0.0035	0.0033	0.004	0.0043
OOS RMSE	0.0039	0.0027	0.0039	0.0038	0.0049	0.0054	0.0053	0.0053

Note: Linear regression models were estimated for eight event windows, including four variables indicating sentiment score, market surprise component, PMI and CCI. Calculations were made for statement release days only between January 2006 and June 2020. Model was obtained using R Software. Significance of variables is assessed using p-value included in the table in round brackets. Significant relations at 5% level were indicated in bold.

Table 3. Support vector regression model results for S&P 500

Predictors	*60-minute before event*	*30-minute before event*	*15-minute after event*	*20-minute after event*	*25-minute after event*	*30-minute after event*	*60-minute after event*	*120-minute after event*
Score	0.003745	0.002521	0.003837	0.003840	0.004597	0.005303	0.004861	0.005727
Market shock	0.003745	0.002521	0.003911	0.003466	0.004603	0.005365	0.005455	0.005243
PMI	0.003750	0.002523	0.003830	0.004176	0.004603	0.005306	0.004881	0.005684
CCI	0.003750	0.002519	0.003829	0.003909	0.004594	0.005307	0.004847	0.005525
Observations (training sample)	96	98	95	95	94	99	94	41
CV MAE	0.0021	0.0014	0.0027	0.0030	0.0031	0.0032	0.0036	0.0040
CV RMSE	0.0033	0.0022	0.0037	0.0039	0.0044	0.0048	0.0048	0.0052
Observations (out-of-sample)	12	12	12	12	12	12	12	12
OOS MAE	0.0017	0.0012	0.0030	0.0031	0.0035	0.0032	0.0041	0.0040
OOS RMSE	0.0035	0.0015	0.0038	0.0038	0.0053	0.0045	0.0061	0.0052
Kernel type	Polynomial	Linear	Linear	Polynomial	Linear	Radial	Linear	Radial
Cost parameter	0.001	0.001	0.001	0.001	0.001	2	0.106	0.316
Sigma parameter	n/a	n/a	n/a	n/a	n/a	0.001	n/a	1

Note: Support vector regression models were estimated for eight event windows, including four variables indicating sentiment score, market surprise component, PMI and CCI. Calculations were made for statement release days only between January 2006 and June 2020. Models were obtained using R Software. 10-fold cross validation was applied. Measures namely CV MAE and CV RMSE were calculated as the average value of *these measures for each data subset. In place of model coefficients as in case of linear regression, the permutation feature importance indicating what is the mean dropout loss of RMSE was used. Seed value of 1234 was used.*

Table 4. Random forest model results for S&P 500

Predictors	*60-minute before event*	*30-minute before event*	*15-minute after event*	*20-minute after event*	*25-minute after event*	*30-minute after event*	*60-minute after event*	*120-minute after event*
Score	0.002879	0.001923	0.002303	0.002640	0.003163	0.003681	0.003613	0.004526
Market shock	0.002757	0.001965	0.003895	0.003639	0.003484	0.004100	0.004226	0.004273
PMI	0.002917	0.001943	0.002410	0.002915	0.003340	0.003828	0.003680	0.004537
CCI	0.002878	0.001913	0.001966	0.002818	0.003367	0.003956	0.003698	0.004476
Observations (training sample)	96	98	95	95	94	99	94	41
CV MAE	0.0025	0.0017	0.0026	0.0029	0.0032	0.0035	0.0037	0.0048
CV RMSE	0.0038	0.0026	0.0034	0.0039	0.0045	0.0050	0.0049	0.0062
Observations (out-of-sample)	12	12	12	12	12	12	12	12
OOS MAE	0.0026	0.0013	0.0042	0.0039	0.0038	0.0040	0.0045	0.0038
OOS RMSE	0.0041	0.0016	0.0055	0.0050	0.0051	0.0050	0.0057	0.0048
mtry parameter	1	1	3	2	1	1	1	1
n-trees parameter	2000	1000	1500	1500	2500	1000	1500	2500

Note: Random forest models were estimated for eight event windows, including four variables indicating sentiment score, market surprise component, PMI and CCI. Calculations were made for statement release days only between January 2006 and June 2020. Models were obtained using R Software. 10-fold cross validation was applied. Measures namely CV MAE and CV RMSE were calculated as the average value of *these measures for each data subset. In place of model coefficients as in case of linear regression, the permutation feature importance indicating what is the mean dropout loss of RMSE was used. Seed value of 1234 was used.*

Table 5. Linear regression model results for EUR/USD

Predictors	*60-minute before event*	*30-minute before event*	*15-minute after event*	*20-minute after event*	*25-minute after event*	*30-minute after event*	*60-minute after event*	*120-minute after event*
(Intercept)	-0.000027	**-0.002470**	-0.000858	0.000725	0.000050	0.000264	0.000265	**0.005703**
	(0.983)	(0.008)	(0.171)	(0.247)	(0.939)	(0.727)	(0.800)	(0.014)
Score	-0.012513	-0.015442	0.015815	0.011847	-0.002414	-0.004145	-0.035225	-0.002119
	(0.760)	(0.600)	(0.444)	(0.566)	(0.911)	(0.869)	(0.309)	(0.978)
Score PFI	0.000911	0.000656	0.000461	0.000460	0.000483	0.000558	0.000779	0.001678
Market shock	**0.001546**	-0.000197	**-0.001123**	**-0.001146**	-0.000477	-0.000028	-0.000546	-0.002124
	(0.016)	(0.662)	(0.001)	(0.000)	(0.153)	(0.943)	(0.302)	(0.069)
Market shock PFI	0.000980	0.000657	0.000536	0.000541	0.000501	0.000558	0.000784	0.001761
PMI	0.000011	**0.000051**	0.000022	-0.000008	0.000006	0.000005	-0.000002	**-0.000115**
	(0.684)	(0.008)	(0.095)	(0.514)	(0.645)	(0.744)	(0.919)	(0.017)
PMI PFI	0.000912	0.000730	0.000482	0.000461	0.000486	0.000560	0.000768	0.001850
CCI	-0.000004	-0.000003	-0.000003	-0.000003	-0.000004	**-0.000005**	-0.000002	0.000006
	(0.316)	(0.358)	(0.132)	(0.169)	(0.090)	(0.043)	(0.558)	(0.437)
CCI PFI	0.000928	0.000665	0.000477	0.000477	0.000516	0.000605	0.000775	0.001692
Observations (training sample)	96	98	95	95	94	99	94	41
MAE	0.0007	0.0005	0.0004	0.0004	0.0004	0.0004	0.0006	0.0012
RMSE	0.0009	0.0007	0.0005	0.0005	0.0005	0.0006	0.0008	0.0017
Observations (out-of-sample)	12	12	12	12	12	12	12	12
OOS MAE	0.0008	0.0006	0.0004	0.0005	0.0005	0.0006	0.0007	0.0011
OOS RMSE	0.0015	0.0011	0.0006	0.0006	0.0008	0.0006	0.0009	0.0020

Note: Linear regression models were estimated for eight event windows, including four variables indicating sentiment score, market surprise component, PMI and CCI. Calculations were made for statement release days only between January 2006 and June 2020. Model was obtained using R Software. Significance of variables is assessed using p-value included in the table in round brackets. Significant relations at 5% level were indicated in bold.

Table 6. Support vector regression model results for EUR/USD

Predictors	*60-minute before event*	*30-minute before event*	*15-minute after event*	*20-minute after event*	*25-minute after event*	*30-minute after event*	*60-minute after event*	*120-minute after event*
Score	0.000902	0.000677	0.000458	0.000456	0.000455	0.000504	0.000785	0.001761
Market shock	0.000980	0.000674	0.000503	0.000433	0.000393	0.000473	0.000804	0.001762
PMI	0.000913	0.000703	0.000440	0.000475	0.000471	0.000590	0.000778	0.001765
CCI	0.000900	0.000668	0.000448	0.000490	0.000465	0.000538	0.000776	0.001761
Observations (training sample)	82	82	83	83	83	83	83	83
CV MAE	0.0006	0.0005	0.0004	0.0005	0.0004	0.0004	0.0007	0.0012
CV RMSE	0.0009	0.0006	0.0005	0.0007	0.0004	0.0005	0.0009	0.0017
Observations (out-of-sample)	12	12	12	12	12	12	12	12
OOS MAE	0.0011	0.0007	0.0004	0.0005	0.0005	0.0004	0.0004	0.0005
OOS RMSE	0.0017	0.0011	0.0005	0.0007	0.0008	0.0008	0.0010	0.0010
Kernel type	Polynomial	Polynomial	Radial	Radial	Radial	Radial	Polynomial	Polynomial
Cost parameter	0.001	0.001	1.579	2	2	0.737	0.001	0.001
Sigma parameter	n/a	n/a	0.013	1	1	1	n/a	n/a

Note: Support vector regression models were estimated for eight event windows, including four variables indicating sentiment score, market surprise component, PMI and CCI. Calculations were made for statement release days only between January 2006 and June 2020. Models were obtained using R Software. 10-fold cross validation was applied. Measures namely CV MAE and CV RMSE were calculated as the average value of *these measures for each data subset. In place of model coefficients as in case of linear regression, the permutation feature importance indicating what is the mean dropout loss of RMSE was used. Seed value of 1234 was used.*

Table 7. Random forest model results for EUR/USD

Predictors	*60-minute before event*	*30-minute before event*	*15-minute after event*	*20-minute after event*	*25-minute after event*	*30-minute after event*	*60-minute after event*	*120-minute after event*
Score	0.000760	0.000499	0.000371	0.000371	0.000371	0.000411	0.000611	0.001336
Market shock	0.000675	0.000465	0.000380	0.000387	0.000369	0.000423	0.000609	0.001294
PMI	0.000708	0.000489	0.000369	0.000369	0.000380	0.000435	0.000582	0.001308
CCI	0.000739	0.000543	0.000392	0.000387	0.000388	0.000443	0.000597	0.001307
Observations (training sample)	82	82	83	83	83	83	83	83
CV MAE	0.0008	0.0005	0.0004	0.0004	0.0004	0.0005	0.0007	0.0014
CV RMSE	0.0010	0.0007	0.0005	0.0005	0.0005	0.0006	0.0009	0.0019
Observations (out-of-sample)	12	12	12	12	12	12	12	12
OOS MAE	0.0011	0.0006	0.0004	0.0005	0.0006	0.0006	0.0008	0.0010
OOS RMSE	0.0017	0.0011	0.0005	0.0006	0.0008	0.0008	0.0010	0.0011
mtry parameter	1	1	1	1	1	1	1	1
n-trees parameter	1000	2500	2500	2500	2500	2500	2500	2500

Note: Random forest models were estimated for eight event windows, including four variables indicating sentiment score, market surprise component, PMI and CCI. Calculations were made for statement release days only between January 2006 and June 2020. Models were obtained using R Software. 10-fold cross validation was applied. Measures namely CV MAE and CV RMSE were calculated as the average value of *these measures for each data subset. In place of model coefficients as in case of linear regression, the permutation feature importance indicating what is the mean dropout loss of RMSE was used. Seed value of 1234 was used.*

6. Conclusions

The primary aim of this paper was to examine the impact of FOMC communication, more specifically FOMC Statements on financial markets. In our analysis the reaction of financial markets was approximated by the returns of S&P 500 Index and EUR/USD spot price. We additionally accounted for market expectations regarding the effective Federal Funds Rate and also Purchasing Managers' Index and Consumer Confidence Index. The latter economic indicators are aimed to control for the attitudes toward economic situation from two different perspectives namely business and individual customers. The analyzed period of time covered years from 2006 to 2020. In this article, we examined three research hypotheses:

1. Information content of FOMC statements significantly influences financial markets reaction directly after the event;
2. Additional control variables, i.e. surprise component, PMI and CB Consumer Confidence allow to better explain the market reaction;
3. Applying non-linear models results in better prediction of market reaction due to explaining more of data variability and taking into account potentially non-linear relationships.

In order to verify the first hypothesis, the tone of the content of FOMC statements was quantified and included as one of the independent variables in the model estimation. The second hypothesis was verified similarly, that is by including additional control variables as features in the modelling process. The third hypothesis was verified by estimating one linear model namely linear regression and two non-linear models namely support vector regression and random forest and comparing estimated models' evaluation metrics.

Firstly, based on a linear regression it was not possible to indicate whether there is significant relation and thus what is the direction of the variable indicating the FOMC statements' sentiment score in case on both S&P 500 Index returns and EUR/USD spot price. However, we managed to investigate the relation using non-linear models and the Explained Artificial Intelligence tool namely Partial Dependency Plot. We found that for EUR/USD spot price, the lower score causes larger reaction of the market while the higher score causes smaller reaction of the market. For S&P 500 Index returns the relation of variable indicating sentiment's score and the model prediction is rather difficult to assess.

Secondly, both linear and non-linear algorithms proved that additional control variables such as market surprise component, Purchasing Managers' Index and Consumer Confidence Index indicate significant (important) relation with both S&P 500 Index returns and EUR/USD spot price. The variable that indicates significant relation the most common with respect to event window is the market surprise component.

Thirdly, it is difficult to unequivocally verify the hypothesis that non-linear models result in better predictions than linear models. We conducted the comparative analysis of forecast errors for linear and non-linear models and the stability analysis of MAE and RMSE measures between the training sample and test sample. The former analysis resulted in the conclusion that the lower average values of out-of-sample MAE and out-of-sample RMSE are observed for the support vector regression model taking into account all event windows. This suggests that especially support vector regression algorithm explains more of data variability than a simple linear regression. The latter analysis resulted in generally lower difference in MAE and RMSE between the training and test sample for linear regression compared to support vector regression and random forest for all considering event windows. Thus, we are able neither to confirm nor to reject the hypothesis that the non-linear models explain more of data variability by taking into account potentially non-linear relationships.

We see several possible extensions on which we will work in the future research. The first aspect regards sentiment extraction. Utilized method, as also indicated before, has several limitations. Possible improvement in results could be obtained by creating a customized dictionary for FOMC statements sentiment extraction based on FOMC communication. Another option is to label the statements and use any classification algorithm. The second aspect regards designing a system which will be fully efficient and enable to obtain real time results. Last but not least, an extension could further regard constructing and back testing an event arbitrage strategy which we would consider the end of the journey with FOMC communication.

References

Aldridge, I. (2013). High-frequency trading: a practical guide to algorithmic strategies and trading systems (Vol. 604). John Wiley & Sons.

Bernanke, B. S., & Kuttner, K. N. (2005). What explains the stock market's reaction to Federal Reserve policy? The Journal of finance, 60(3), 1221-1257.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In Proceedings of the fifth annual workshop on Computational learning theory (pp. 144-152).

Boukus, E., & Rosenberg, J. V. (2006). The information content of FOMC minutes. Available at SSRN 922312.

Cannon, S. (2015). Sentiment of the FOMC: Unscripted. Economic Review-Federal Reserve Bank of Kansas City, 5.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

Cieslak, A., Morse, A., & Vissing-Jorgensen, A. (2019). Stock returns over the FOMC cycle. The Journal of Finance, 74(5), 2201-2248.

Farka, M., & Fleissig, A. R. (2013). The impact of FOMC statements on the volatility of asset prices. Applied Economics, 45(10), 1287-1301.

Gidofalvi, G., & Elkan, C. (2001). Using news articles to predict stock price movements. Department of Computer Science and Engineering, University of California, San Diego.

Guyon, I., Boser, B., & Vapnik, V. (1993). Automatic capacity tuning of very large VC-dimension classifiers. In Advances in neural information processing systems (pp. 147-155).

Hayo, B., Kutan, A., & Neuenkirch, M. (2008). Communicating With Many Tongues: The Impact of FOMC Members Speeches on the US Financial Market's Returns and Volatility. In Paper to be presented on the 40th Money, Macro and Finance Research Group annual conference, September (pp. 10-12).

Jegadeesh, N., & Wu, D. A. (2017). Deciphering Fedspeak: The information content of FOMC meetings. Available at SSRN 2939937.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. Journal of monetary economics, 47(3), 523-544.

Schumaker, R. P., Zhang, Y., Huang, C. N., & Chen, H. (2012). Evaluating sentiment in financial news articles. Decision Support Systems, 53(3), 458-464.

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. Journal of Accounting Research, 54(4), 1187-1230.

Lucca, D. O., & Trebbi, F. (2009). Measuring central bank communication: an automated approach with application to FOMC statements (No. w15367). National Bureau of Economic Research.

Mazis, P., & Tsekrekos, A. (2017). Latent semantic analysis of the FOMC statements. Review of Accounting and Finance, 16(2), 179–217. doi:10.1108/raf-10-2015-0149.

Rosa, C. (2012). How "Unconventional" are Large-Scale Asset Purchases? The Impact of Monetary Policy on Asset Prices. SSRN Electronic Journal. doi:10.2139/ssrn.2053640.

Vapnik, V., Golowich, S. E., & Smola, A. J. (1997). Support vector method for function approximation, regression estimation and signal processing. In Advances in neural information processing systems (pp. 281-287).

Zadeh, R. B., & Zollmann, A. (2009). Predicting Market-Volatility from Federal Reserve Board Meeting Minutes NLP for Finance.

Appendix A

January 30, 2019

"In determining the timing and size of future adjustments to the target range for the federal funds rate, the Committee will assess realized and expected economic conditions relative to its maximum employment objective and its symmetric 2 percent inflation objective. This assessment will take into account a wide range of information, including measures of labor market conditions, indicators of inflation pressures and inflation expectations, and readings on financial and international developments."

Appendix B

January 31, 2006

"The Federal Open Market Committee decided today to raise its target for the federal funds rate by 25 basis points to 4-1/2 percent. Although recent economic data have been uneven, the expansion in economic activity appears solid. Core inflation has stayed relatively low in recent months and longer-term inflation expectations remain contained. Nevertheless, possible increases in resource utilization as well as elevated energy prices have the potential to add to inflation pressures. The Committee judges that some further policy firming may be needed to keep the risks to the attainment of both sustainable economic growth and price stability roughly in balance. In any event, the Committee will respond to changes in economic prospects as needed to foster these objectives."

December 17, 2014

"Information received since the Federal Open Market Committee met in October suggests that economic activity is expanding at a moderate pace. Labor market conditions improved further, with solid job gains and a lower unemployment rate. On balance, a range of labor market indicators suggests that underutilization of labor resources continues to diminish. Household spending is rising moderately and business fixed investment is advancing, while the recovery in the housing sector remains slow. Inflation has continued to run below the Committee's longer-run objective, partly reflecting declines in energy prices. Market-based measures of inflation compensation have declined somewhat further; survey-based measures of longer-term inflation expectations have remained stable.

Consistent with its statutory mandate, the Committee seeks to foster maximum employment and price stability. The Committee expects that, with appropriate policy accommodation, economic activity will expand at a moderate pace, with labor market indicators moving toward levels the Committee judges consistent with its dual mandate. The Committee sees the risks to the outlook for economic activity and the labor market as nearly balanced. The Committee expects inflation to rise gradually toward 2 percent as the labor market improves further and the transitory effects of lower energy prices and other factors dissipate. The Committee continues to monitor inflation developments closely.

To support continued progress toward maximum employment and price stability, the Committee today reaffirmed its view that the current 0 to 1/4 percent target range for the federal funds rate remains appropriate. In determining how long to maintain this target range, the Committee will assess progress--both realized and expected--toward its objectives of maximum employment and 2 percent inflation. This assessment will take into account a wide range of information, including measures of labor market conditions, indicators of inflation pressures and inflation expectations, and readings on financial developments. Based on its current assessment, the Committee judges that it can be patient in beginning to normalize the stance of monetary policy. The Committee sees this guidance as consistent with its previous statement that it likely will be appropriate to maintain the 0 to 1/4 percent target range for the federal funds rate for a considerable time following the end of its asset purchase program in October, especially if projected inflation continues to run below the Committee's 2 percent longer-run goal, and provided that longer-term inflation expectations remain well anchored. However, if incoming information indicates faster progress toward the Committee's employment and inflation objectives than the Committee now expects, then increases in the target range for the federal funds rate are likely to occur sooner than currently anticipated. Conversely, if progress proves slower than expected, then increases in the target range are likely to occur later than currently anticipated.

The Committee is maintaining its existing policy of reinvesting principal payments from its holdings of agency debt and agency mortgage-backed securities in agency mortgage-backed securities and of rolling over maturing Treasury securities at auction. This policy, by keeping the Committee's holdings of longer-term securities at sizable levels, should help maintain accommodative financial conditions.

When the Committee decides to begin to remove policy accommodation, it will take a balanced approach consistent with its longer-run goals of maximum employment and inflation of 2 percent. The Committee currently anticipates that, even after employment and inflation are near mandate-consistent levels, economic conditions may, for some time, warrant keeping the target federal funds rate below levels the Committee views as normal in the longer run."

July 29, 2020

"The Federal Reserve is committed to using its full range of tools to support the U.S. economy in this challenging time, thereby promoting its maximum employment and price stability goals.

The coronavirus outbreak is causing tremendous human and economic hardship across the United States and around the world. Following sharp declines, economic activity and employment have picked up somewhat in recent months but remain well below their levels at the beginning of the year. Weaker demand and significantly lower oil prices are holding down consumer price inflation. Overall financial conditions have improved in recent months, in part reflecting policy measures to support the economy and the flow of credit to U.S. households and businesses.

The path of the economy will depend significantly on the course of the virus. The ongoing public health crisis will weigh heavily on economic activity, employment, and inflation in the near term, and poses considerable risks to the economic outlook over the medium term. In light of these developments, the Committee decided to maintain the target range for the federal funds rate at 0 to 1/4 percent. The Committee expects to maintain this target range until it is confident that the economy has weathered recent events and is on track to achieve its maximum employment and price stability goals.

The Committee will continue to monitor the implications of incoming information for the economic outlook, including information related to public health, as well as global developments and muted inflation pressures, and will use its tools and act as appropriate to support the economy. In determining the timing and size of future adjustments to the stance of monetary policy, the Committee will assess realized and expected economic conditions relative to its maximum employment objective and its symmetric 2 percent inflation objective. This assessment will take into account a wide range of information, including measures of labor market conditions, indicators of inflation pressures and inflation expectations, and readings on financial and international developments.

To support the flow of credit to households and businesses, over coming months the Federal Reserve will increase its holdings of Treasury securities and agency residential and commercial mortgage-backed securities at least at the current pace to sustain smooth market functioning, thereby fostering effective transmission of monetary policy to broader financial conditions. In addition, the Open Market Desk will continue to offer large-scale overnight and term repurchase agreement operations. The Committee will closely monitor developments and is prepared to adjust its plans as appropriate."

Appendix C

Before text processing:

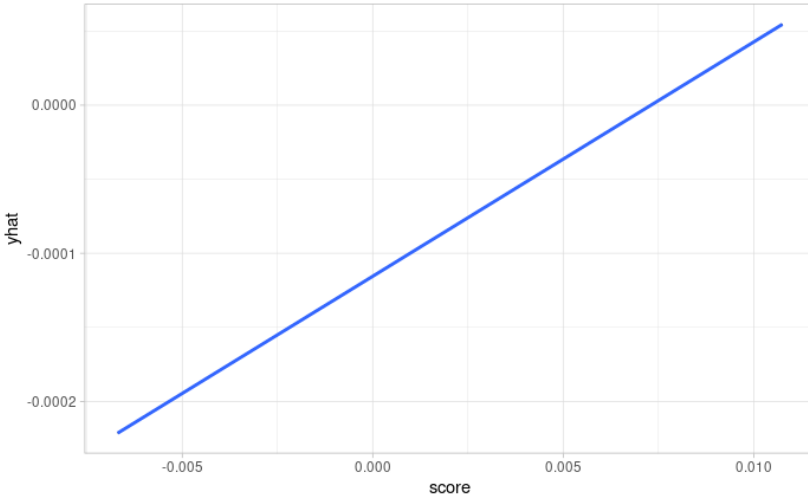
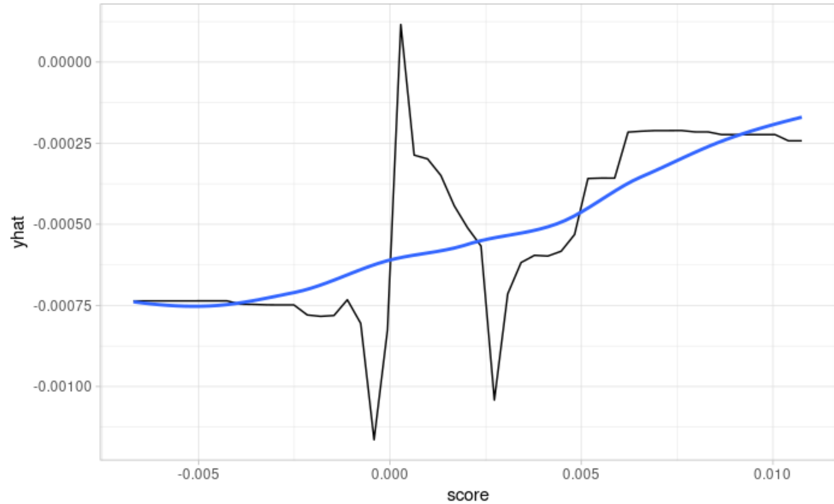
The Federal Open Market Committee decided today to raise its target for the federal funds rate by 25 basis points to 4-1/2 percent. Although recent economic data have been uneven, the expansion in economic activity appears solid. Core inflation has stayed relatively low in recent months and longer-term inflation expectations remain contained. Nevertheless, possible increases in resource utilization as well as elevated energy prices have the potential to add to inflation pressures. The Committee judges that some further policy firming may be needed to keep the risks to the attainment of both sustainable economic growth and price stability roughly in balance. In any event, the Committee will respond to changes in economic prospects as needed to foster these objectives.

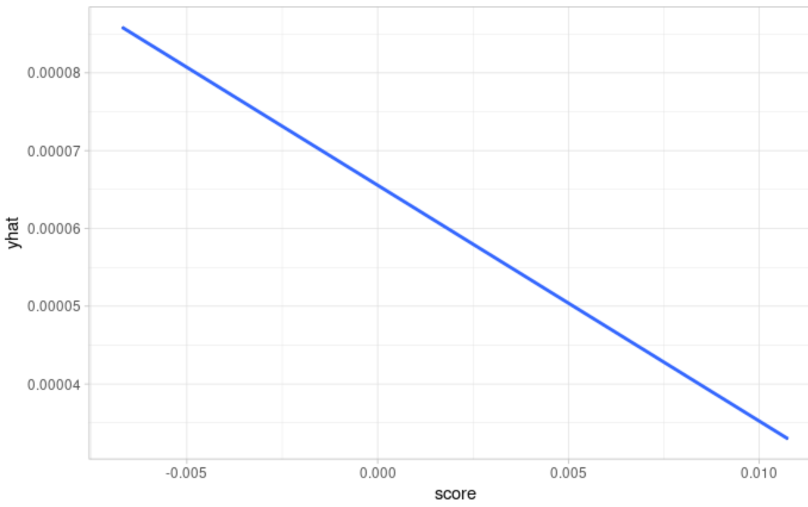
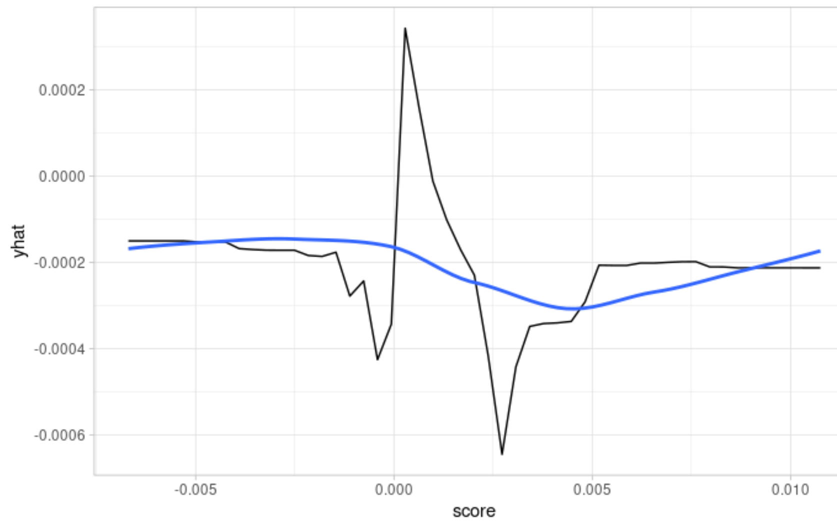
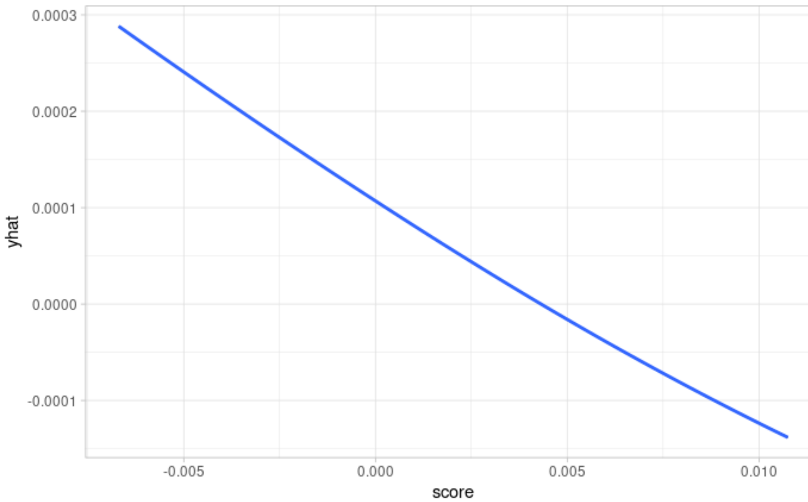
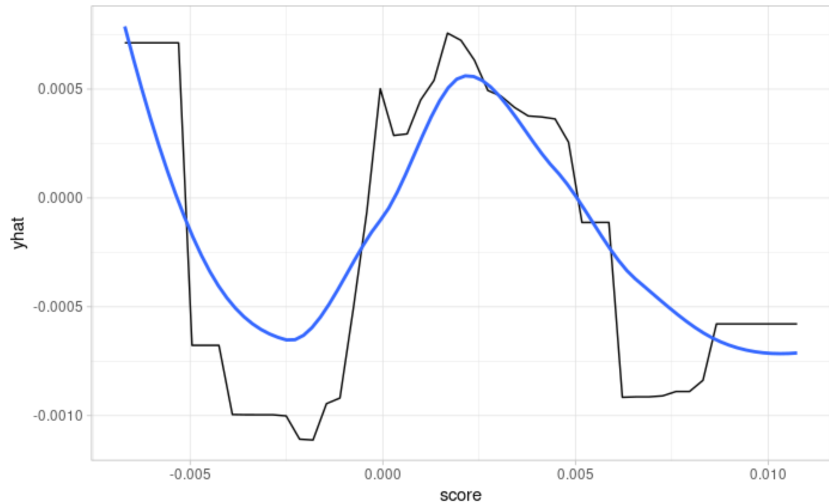
After text processing:

decide today raise target fund rate basis point percent recent economic datum uneven expansion economic activity appear solid core inflation stay low recent month longerterm inflation expectation remain contain increase resource utilization elevate energy price potential add inflation pressure judge policy firm need risk attainment sustainable economic growth price stability roughly balance event respond economic prospect need foster objective

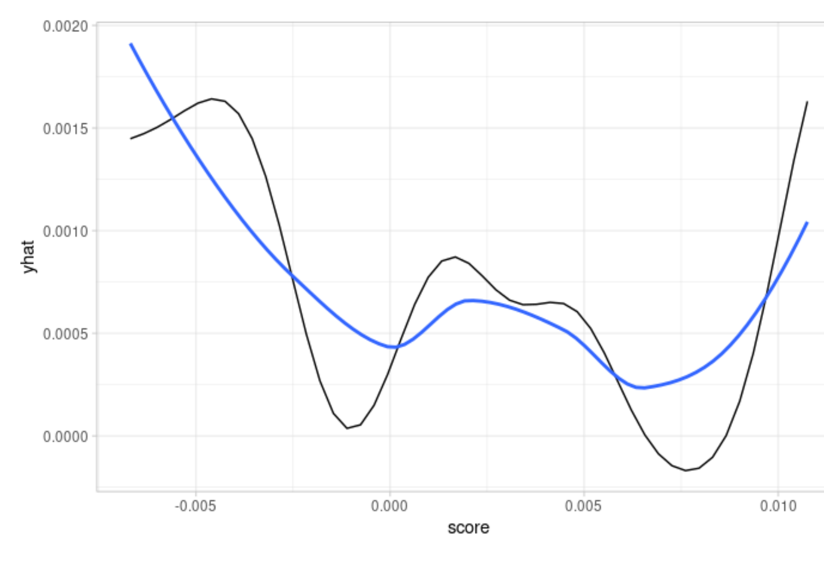
Appendix D

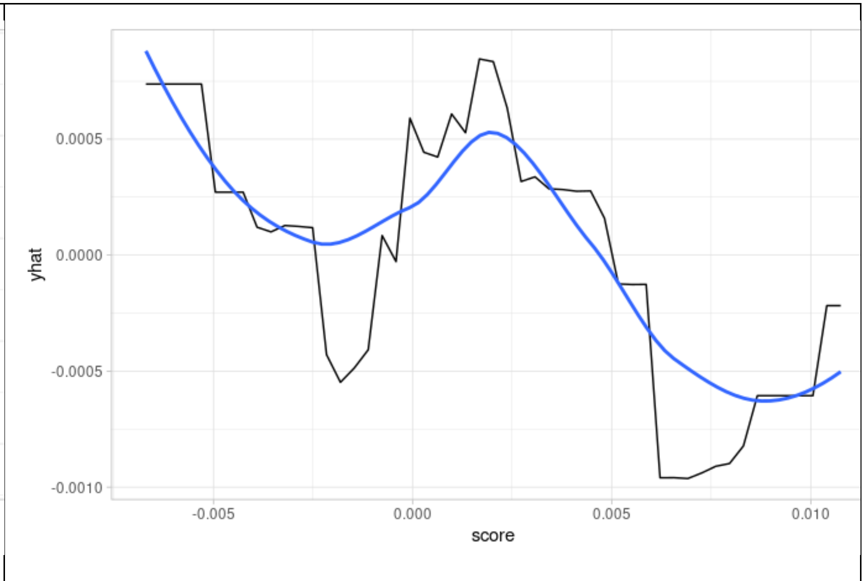
Table D1: Partial dependency plots for variable indicating sentiment score for S&P 500 Index returns model estimation

Event window	Support vector regression	Random forest
60-minute before		

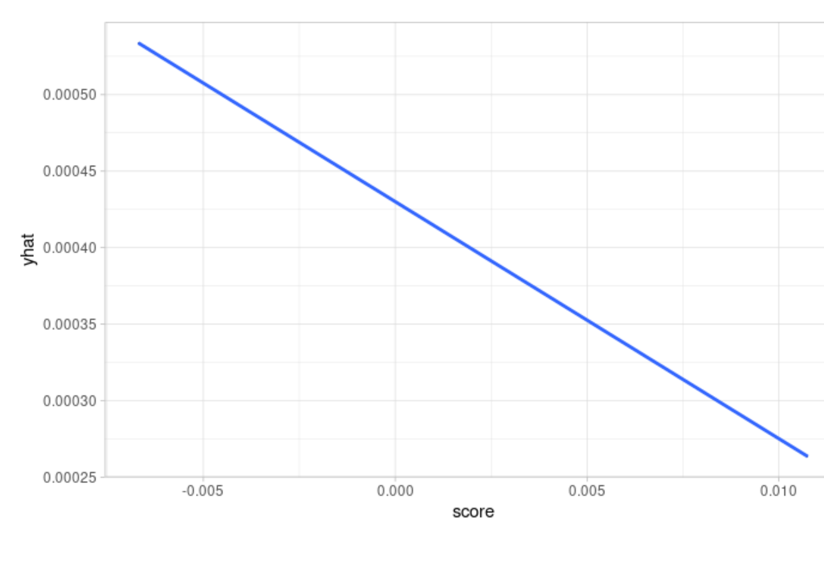
30-minute before		
15-minute after		

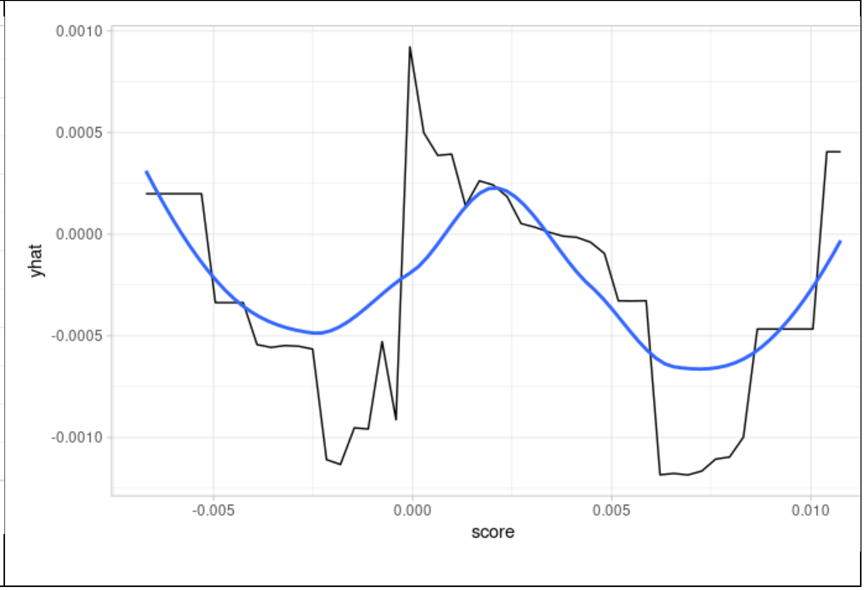
20-minute after

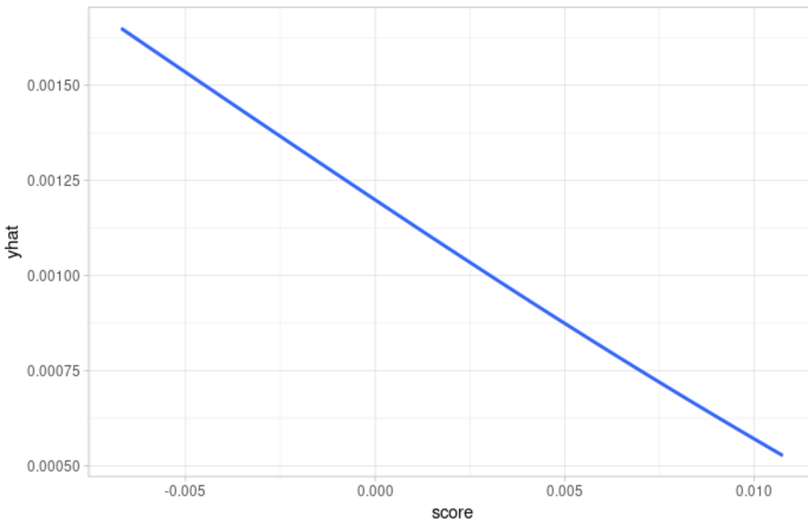
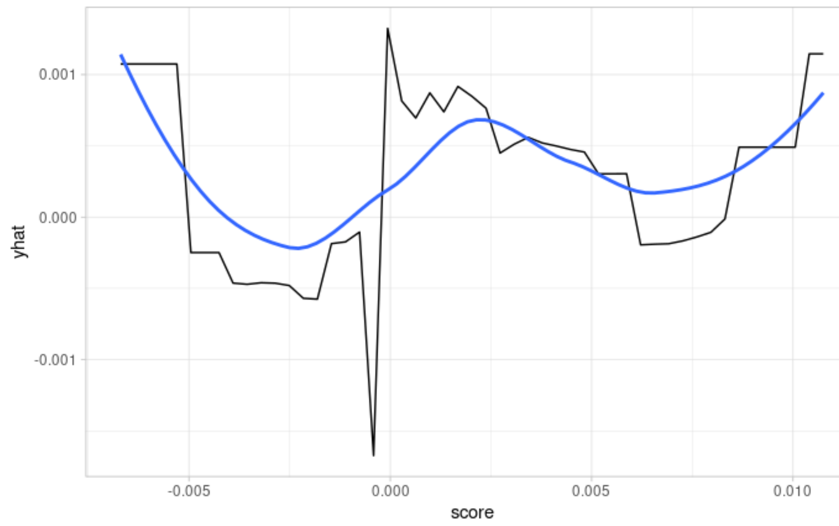
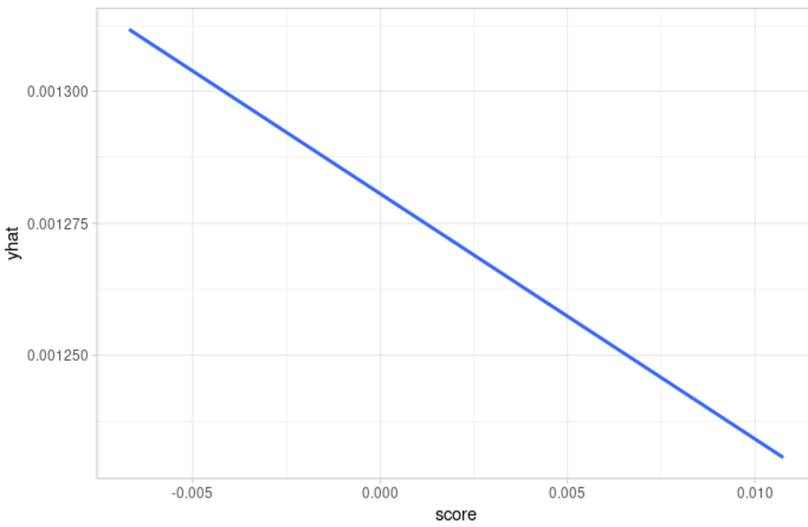
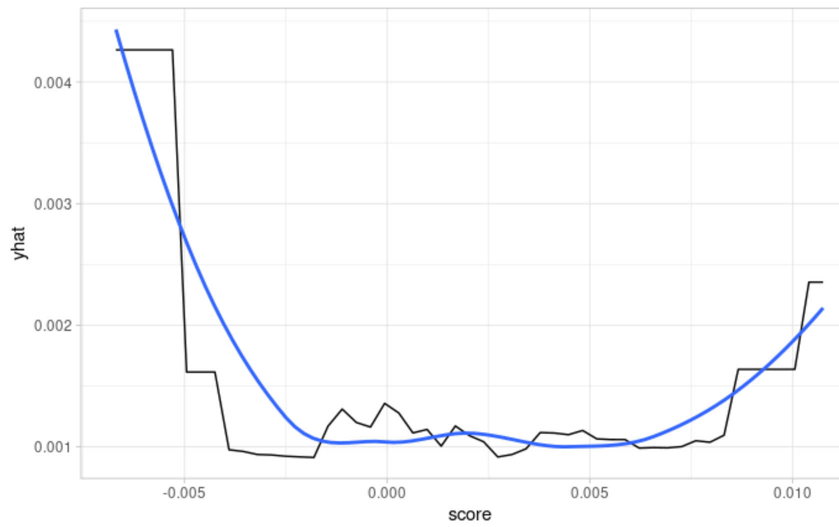


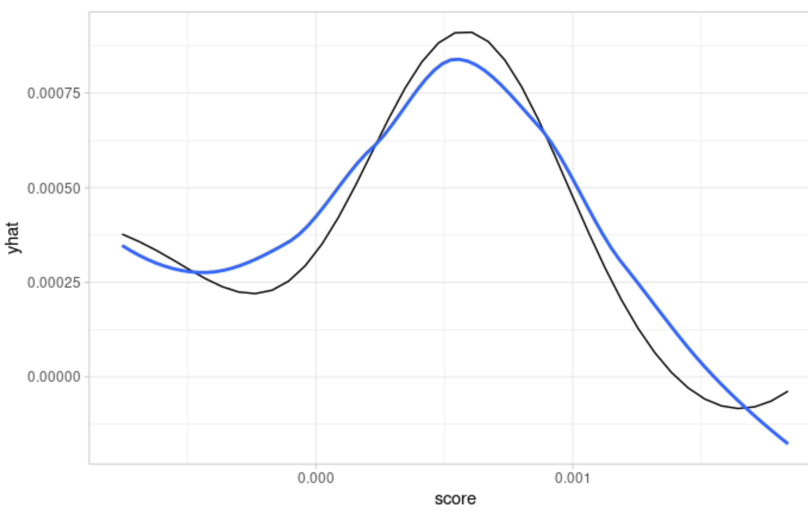
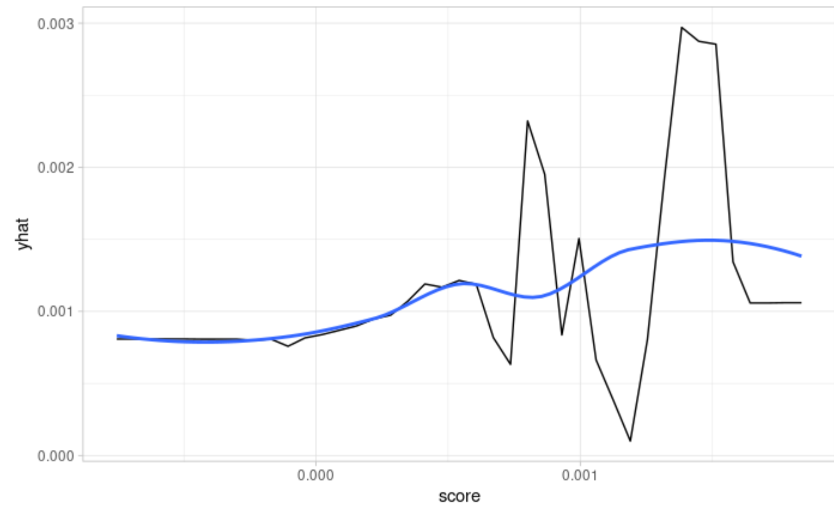


25-minute after



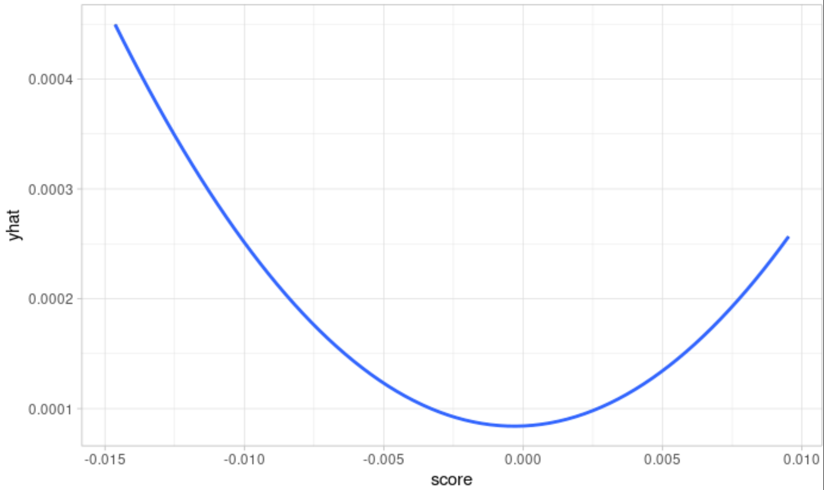
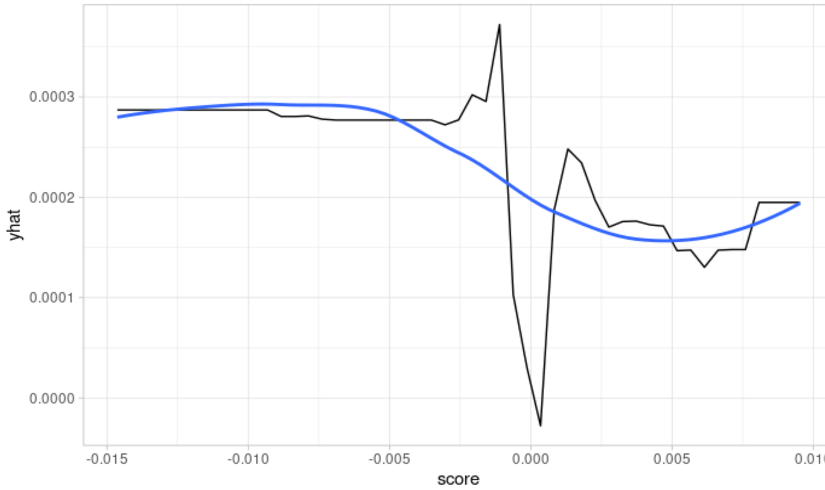


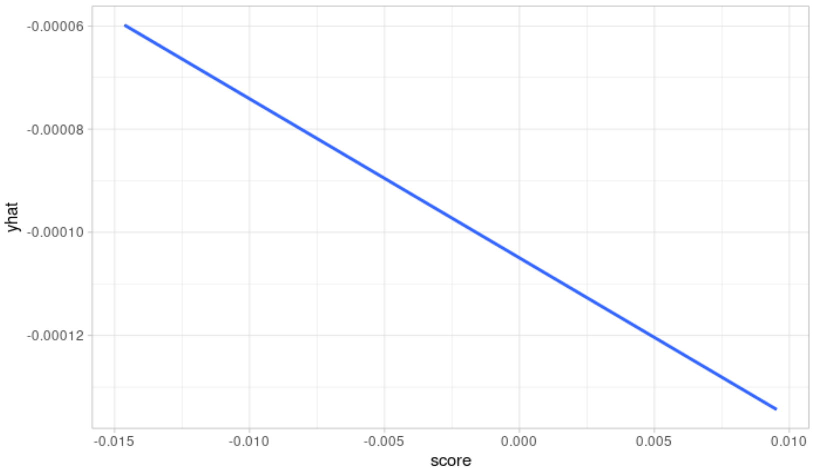
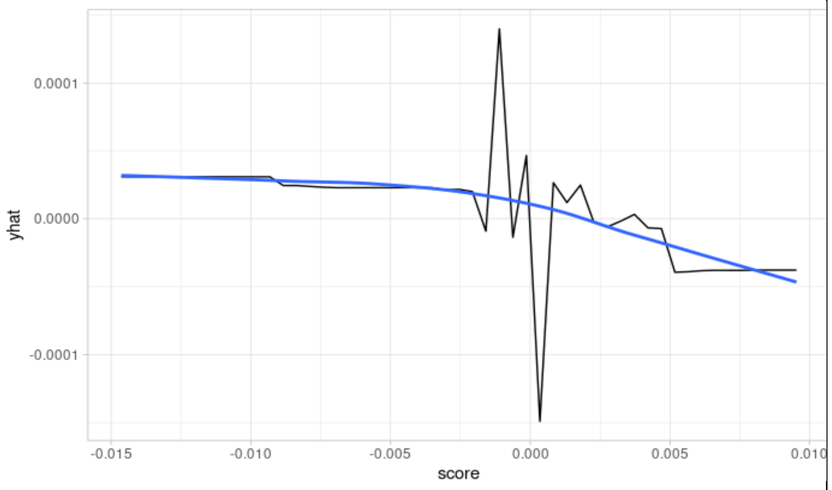
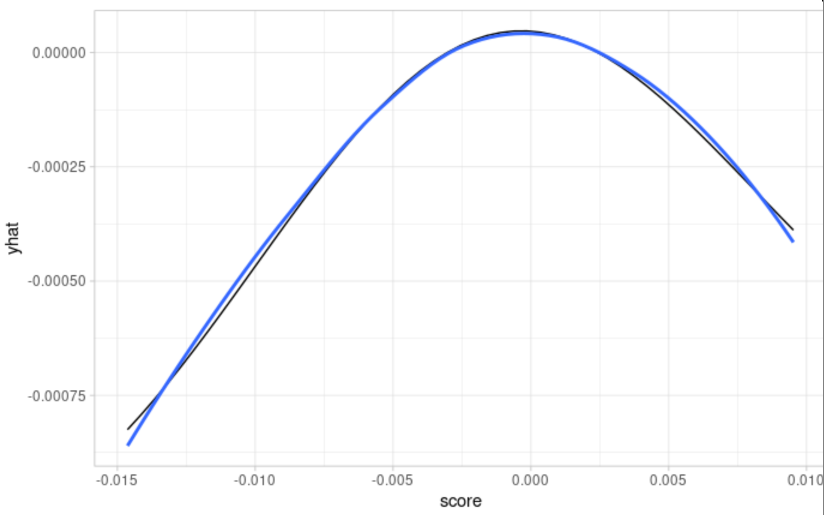
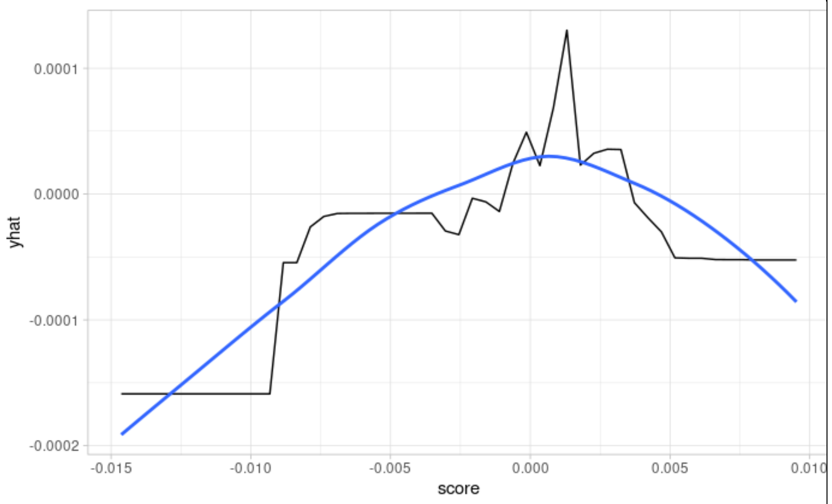
30-minute after		
60-minute after		

120-minute after		

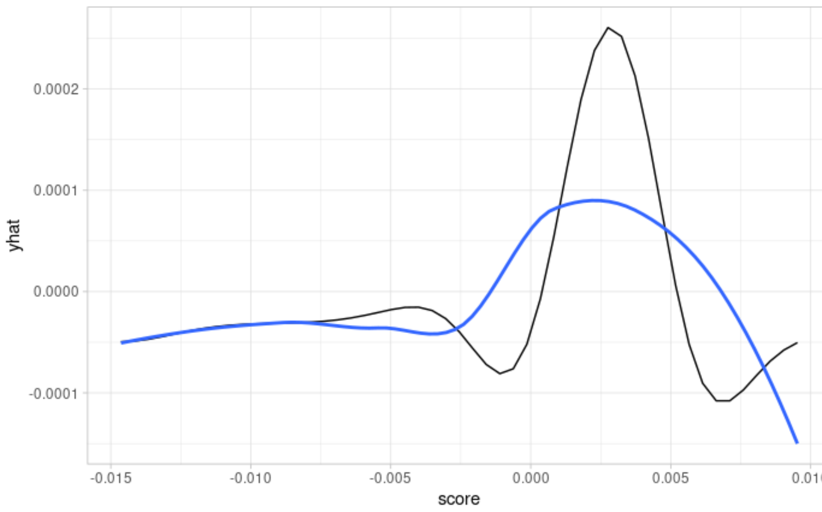
Notes: Partial dependency plots were obtained based on the estimated models which results were presented in the main article text. Plots present the dependence between target variable (indicated at the plot as yhat) and variable indicating FOMC statements' sentiment score. The analysis concerns support vector regression and random forest models. In case of support vector regression, straight line at the plot means the linear kernel was used for model estimation.

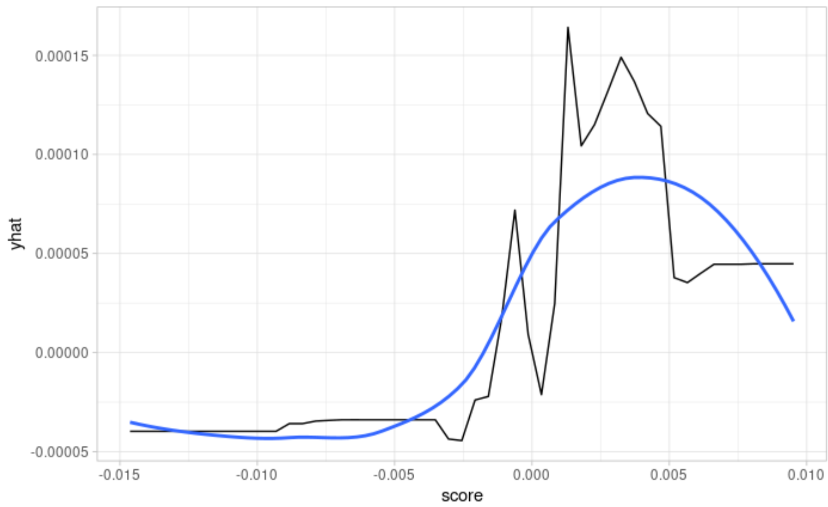
Table D2: Partial dependency plots for variable indicating sentiment score for EUR/USD spot price model estimation

Event window	Support vector regression	Random forest
60-minute before		

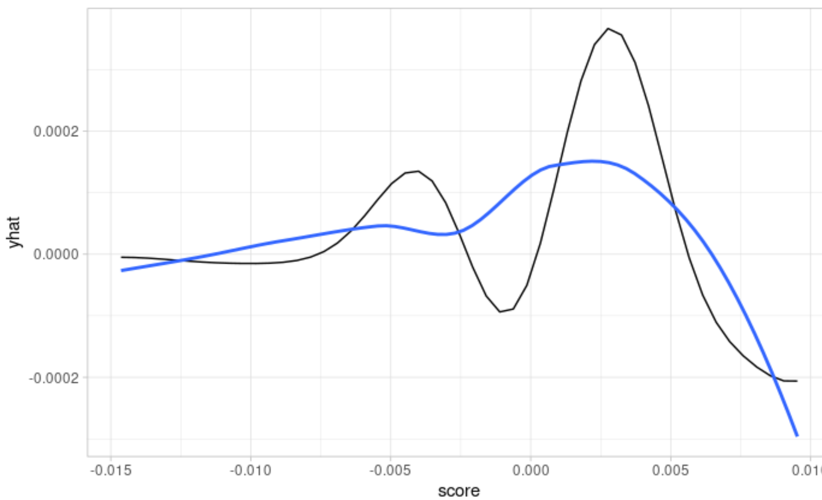
<table>
<tr>
<td>30-minute before</td>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>15-minute after</td>
<td>
</td>
<td>
</td>
</tr>
</table>

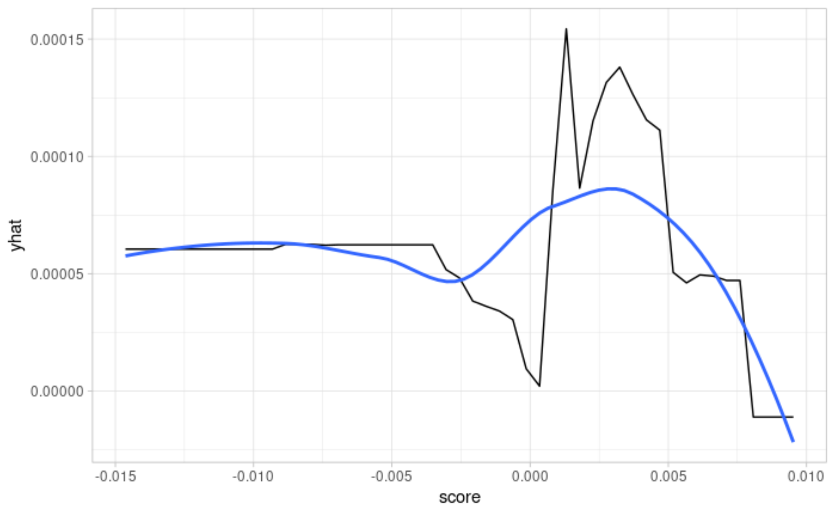
20-minute after



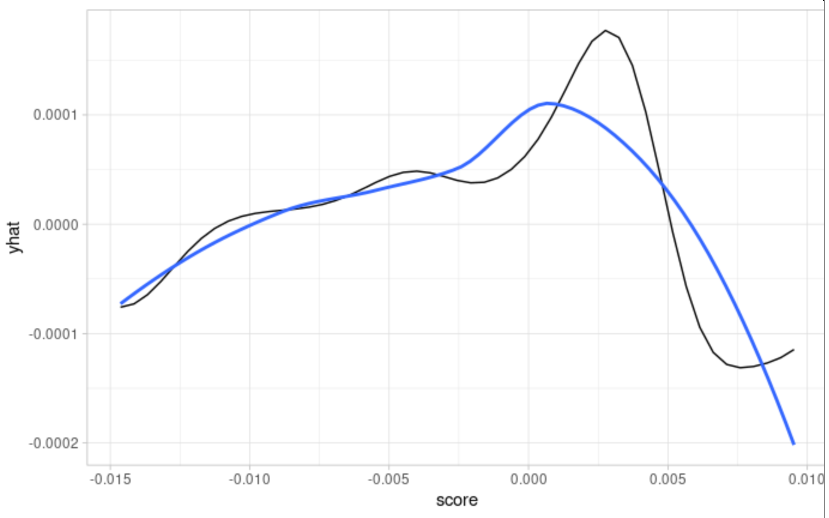


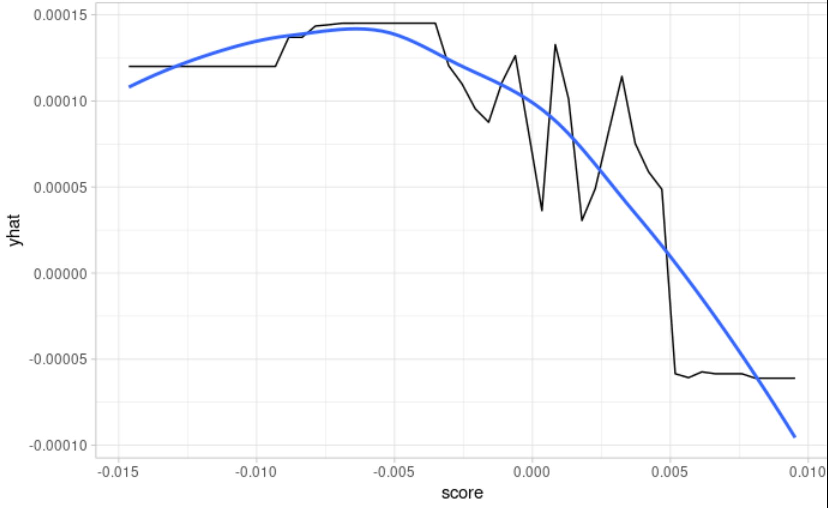
25-minute after



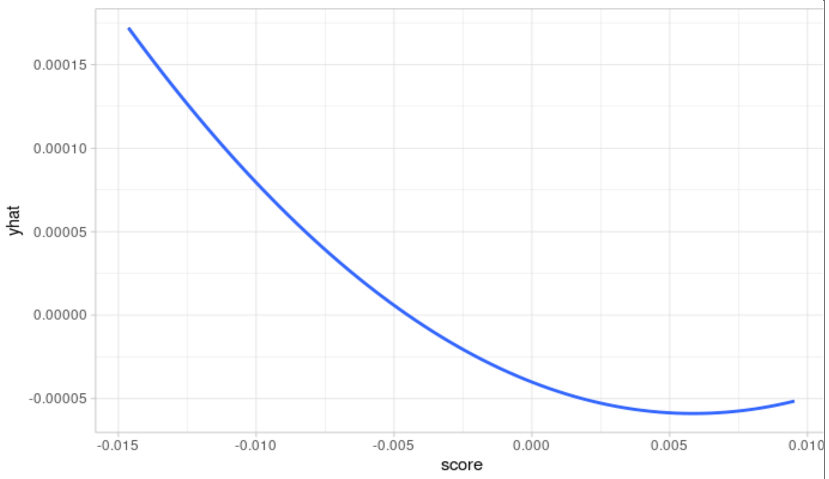


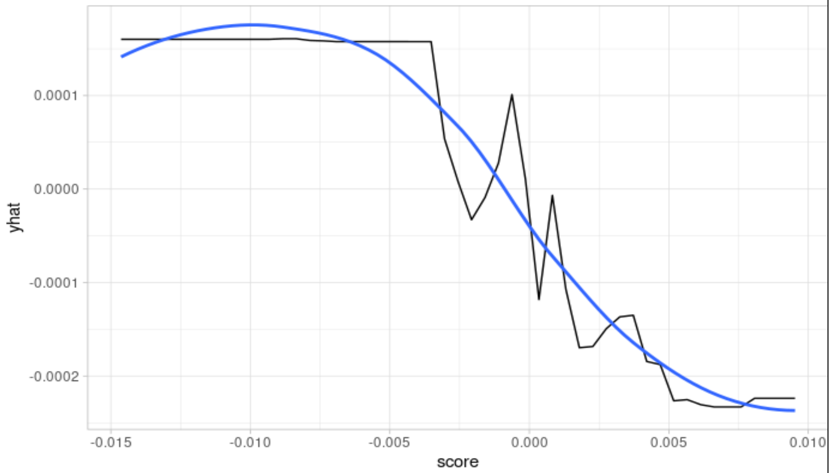
30-minute after



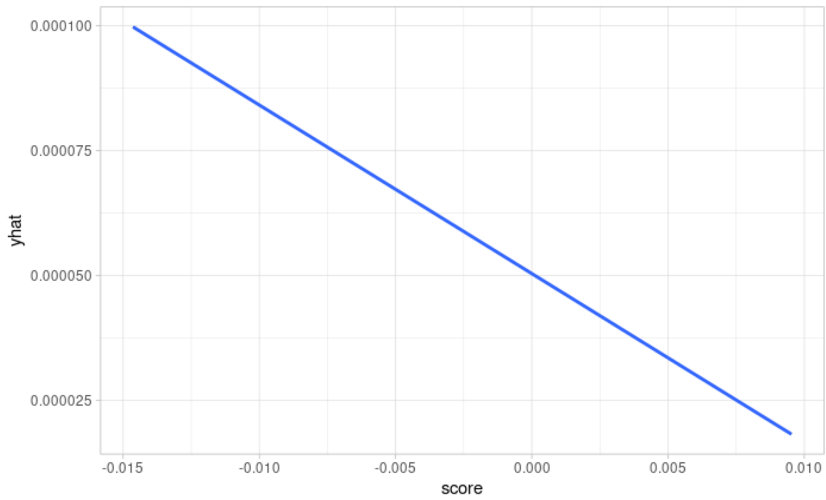


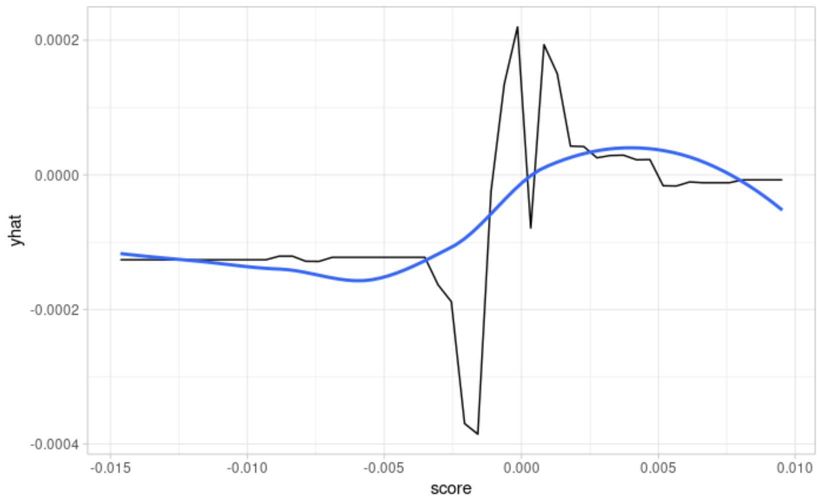
60-minute after





120-minute after





Notes: Partial dependency plots were obtained based on the estimated models which results were presented in the main article text. Plots present the dependence between target variable (indicated at the plot as yhat) and variable indicating FOMC statements' sentiment score. The analysis concerns support vector regression and random forest models. In case of support vector regression, straight line at the plot means the linear kernel was used for model estimation.

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl