



Munich Personal RePEc Archive

## **Central Bank Communication: Information and Policy shocks**

Ostapenko, Nataliia

University of Tartu

22 May 2020

Online at <https://mpra.ub.uni-muenchen.de/101278/>  
MPRA Paper No. 101278, posted 29 Jun 2020 09:48 UTC

# Central Bank Communication: Information and Policy shocks

Nataliia Ostapenko \*

June 21, 2020

## Abstract

The study proposes an alternative way to decompose Federal Reserve (Fed) information shocks from monetary policy shocks by employing a textual analysis to Federal Open Market Committee (FOMC) statements. I decompose Fed statements into economic topics using Latent Dirichlet Allocation (LDA). The model was trained on the business section from major US newspapers. After decomposing surprises in Fed futures into a part that is explained by topics from the Fed statements and that is not explained, the study employs these purged series as proxies for monetary policy and Fed information shocks. The results show that, compared to surprises in 3-month federal funds futures, a policy shock identified in this study has a more negative effect on GDP and a more prolonged negative effect on inflation. In the short-run it causes S&P500 to decline and the Fed to raise its interest rate. Identified Fed information shock affects the macroeconomy as the standard news shock: it has positive long-run effects on S&P500, interest rates, and real GDP, whereas it has a negative short-run effect on inflation. Moreover, the Fed information shock reduces credit costs.

**Keywords:** FOMC, statements, Latent Dirichlet Allocation, monetary policy, information, shocks

**JEL Classification:** E52, E31, E00

---

\*Department of Economics and Business Administration, University of Tartu. Contact: nataliia.ostapenko@gmail.com

# 1 Introduction

How monetary policy affects the economy? For answering this question one must find a good measure for monetary policy shocks. That is because the Fed reacts to macroeconomic indicators and shocks should be orthogonal to this reaction. The main empirical strategies lie in purging a monetary policy instrument from the reaction function (Romer & Romer (2004)) or employing high-frequency identification (Gertler & Karadi (2015)). But the recent studies pointed out that the information effect of central bank communication might invalidate even high-frequency identification (Steinsson (2019), Jarocinski & Karadi (2020), Hansen & McMahon (2016) among others).

The main concern with high-frequency identification lies in the fact that FOMC might possess insider information (Romer & Romer (2000)). As a consequence, FOMC statements might release this private information to the public, and the reaction in a narrow window might contain a response to this additional information instead of a response to unexpected monetary policy action by the Fed. Therefore, a response in 3-month federal funds futures would not be a causal consequence of a monetary policy action itself. In line with that, Miranda-Agrippino & Ricco (2014) found that shocks identified by purging can be predicted from the macroeconomic indicators (from *Federal Reserve Economic Data* (2019)), while shocks identified by the high-frequency strategy are predicted from the *Greenbook Historical and Forecast Data* (2019) projections.

This paper provides original empirical evidence about the information contained in FOMC statements. A distinction of what type of information is important to “policy surprises” allows decomposing these surprises into information and policy effects.

I use FOMC statements as the main data source for 1994–2016 since the Fed started to release statements from 1994. I use Latent Dirichlet Allocation (LDA) pre-trained on the business sections of main US newspapers for content extraction from the FOMC statements. Afterwards, I adopt a lexical-based approach to assign the tone to each sentence from the FOMC statements. The lexical approach counts the proportions of positive/negative and uncertain words in each sentence.

These topic time series are employed to identify the types of information that are important for surprise changes in 3-month federal funds futures on FOMC statement release dates. I use Bayesian Lasso regression for this purpose. Furthermore, the study decom-

poses federal funds future surprises on FOMC dates into information and information-free shocks. These shocks are employed in Structural Vector Autoregressions (SVARs), which lets us disentangle the Fed information effects from a pure policy shock. For this purpose, I use the data from Jarocinski & Karadi (2020) on 3-month federal funds future and S&P500 surprises in a narrow window around FOMC announcement, as well as the main macroeconomic indicators employed in VAR by Jarocinski & Karadi (2020) to make the findings comparable.

The main results are as follows. The important topics are about the economy, monetary policy, credit, investment, company news and deals. After decomposing surprises in Fed futures into the part that is explained by topics from Fed statements and that is not explained, the study uses these purged series as proxies for monetary policy and Fed information shocks. The results show that a policy shock has a more negative effect on GDP and more prolonged negative effect on inflation compared to the baseline results. In the short-run it causes S&P500 to decline and the Fed to raise its interest rate. The Fed information shock has a positive long-run effect on S&P500, as well as on the interest rate, on real GDP, and a negative short-run effect on inflation. Moreover, it reduces the costs of credit.

The findings add to the results of Jarocinski & Karadi (2020), who employed sign restrictions to identify monetary policy and information effects of the Central Bank. First, the baseline results using Jarocinski & Karadi (2020) original variables and recursive identification are of smaller magnitude than in Jarocinski & Karadi (2020) and the response of S&P500 is not robust. The reason is the different studied periods. Second, signs of the information effect based on decomposition employed in this paper are completely in line with the results of Jarocinski & Karadi (2020). Third, the effect of policy surprise shocks are also in line with the main findings of Jarocinski & Karadi (2020): the effect on interest rate is less persistent and on inflation is more persistent.

Moreover, the study extends the findings of Romer & Romer (2000) about asymmetric information between the Federal Reserve and the public. My findings show that there is additional information contained in FOMC statements as well, not just in monetary policy actions themselves.

The information shock, identified in this study, has an expansionary effect on the economy as in Steinsson (2019), who showed that a contractionary monetary policy shock from high-frequency identification has an expansionary effect on output growth expectations.

The results are in line with the results of Hubert & Labondance (2017), who found that sentiment affects private interest rate expectations, inflation and industrial production beyond monetary shocks. Hubert (2019) found that contractionary monetary policy has negative effects on inflation expectations and stock prices only and only if associated with positive economic news. Iglesias et al. (2017) found that neither communication has particularly significant effects on inflation nor real economic activity, whereas this study finds that communication affects inflation and economic activity.

Last but not least, this study complements the recent literature in the way of decomposing FOMC statements into topic time series with sentiments. To the author's best knowledge this is the first study that employs pre-trained LDA model for decomposing the sentences from FOMC statements into economic topic time series. Hansen et al. (2019) used Bank of England Inflation reports and treated each paragraph as a document in LDA. Similarly, Hansen & McMahon (2016) trained the LDA model on sentences from FOMC statements. Afterwards, the authors assigned the tone to each topic. My approach differs from the above-mentioned in that the LDA model was trained on the US newspapers, which lets us obtain more distinguishable topics.

The remainder of the paper proceeds as follows. Section 2 describes the data and methodology for topic extraction from FOMC statements. Section 3 discusses the information content of Fed communication. Section 4 presents the results of purging surprises in 3-month federal funds futures into explained and unexplained parts from FOMC statements. Section 5 concludes.

## 2 Methodology

The Federal Open Market Committee (FOMC) holds eight regularly scheduled meetings during the year and additional meetings as needed. On these meetings Federal Open Market Committee decides on interest rate changes for adjusting inflation. Beginning with the 1994 meetings, the FOMC Secretariat started to release FOMC statements to the public (*Federal Open Market Committee: Transcripts and other historical materials* (2019)). Federal Open Market Committee statements for 1994-2020 were downloaded from the Fed webpage.

The standard high-frequency identification strategy employs a narrow window (30 minutes) to detect surprise changes in 3-month federal funds futures around FOMC announce-

ments. The main concern with this identification strategy lies in the fact that FOMC might possess insider information (Romer & Romer (2000)), and FOMC announcements, therefore, might release new information to the public. The reaction in a narrow window might contain a response to this additional information instead of a response to unexpected monetary policy action. This might invalidate the interpretation of the results based on high-frequency identification since it is not possible to distinguish the effect of monetary policy shocks from information shocks.

Following the logic, Miranda-Agrippino & Ricco (2014) found that Gertler & Karadi (2015) surprises are predictable from *Greenbook Historical and Forecast Data* (2019) projections and *Federal Reserve Economic Data* (2019) factors. The authors purged the shock series with respect to its own lags and Greenbook information (as in Romer & Romer (2004)). But these surprises might just convey the Fed information effect. Contrary to that, I purge shock series with respect to topics from FOMC statements. These topics and the tone of the Fed should capture the Fed information effect and allows to disentangle pure monetary policy shocks from information shocks.

I use the data from Jarocinski & Karadi (2020), who decomposed surprise and information shocks from surprises in 3-month federal funds futures and stock prices around FOMC announcement using sign restrictions. Alternatively, to purge the surprises in federal funds futures from the information effect I decompose these surprises to predictable ( $\widehat{ffr\_hft}$ ) and non-predictable ( $\epsilon_t$ ) components from the regression (1):

$$ffr\_hft = \beta_0 + \sum_{i=1}^K \beta_i info_t^i + \epsilon_t \quad (1)$$

where the dependent variable is a “policy shocks”,  $Info_t^i$  is the information contained in FOMC announcements (described below).

To train a model for the topic extraction (details are presented below) I use the Nexis Uni database, from where I extracted daily business news from The New York Times 1980–2019, The Washington Post 1981–2019, The Los Angeles Times 1985–2019, The Chicago Tribune 1985–2019. The New York Times is the second-largest in circulation and the largest circulating metropolitan newspapers with a weekly circulation of 2.1 million. It is also ranked the 18th in terms of world circulation. The Los Angeles Times is the fourth-largest US newspaper by circulation, The Chicago Tribune is the sixth- and The Washington Post is the seventh-largest US newspaper by circulation. The total timespan is 1980:M6–2019:M7.

Following Shapiro et al. (2017) I filtered out the news that does not contain one of the following words: said, says, told, stated, wrote, reported. After imposing these criteria, the data pull yielded approximately 416,000 articles.

Following Larsen & Thorsrud (2019), I employ Latent Dirichlet Allocation (LDA) (introduced by Blei et al. (2003)) for topic extraction. The LDA is a probabilistic graphical model that is based on the bag-of-words assumption, that is the word order does not matter. If one mixes words in an article and employs the LDA that leads to the same results as without mixing. For extracting news topics with Latent Dirichlet Allocation standard text processing steps are employed:

- Words from a stoplist are excluded. This list contains common words that contribute little meaning to the documents, like prepositions, conjunctions, and pronouns.
- Words are reduced to their word root form. Example: economy, economic, economical, economics, economise are reduced to the root form econom.
- Rare and frequent words are removed
- Vocabulary consists of 57,990 unique words.

LDA is a mixed-membership directed probabilistic graphical model for a text corpus. The generative process for a document collection  $D$  under the LDA model is as follows (Darling (2011)):

1. For each topic  $k = 1, \dots, K$  ( $K$  is the total number of latent topics):
  - A discrete probability distribution over a fixed vocabulary that represents the  $k^{th}$  topic distribution,  $\varphi_k \sim Dirichlet(\beta)$ <sup>1</sup>
2. For each document  $d \in D$  ( $D$  is the total number of documents):
  - A document-specific distribution over the available topics (per-document topic proportion),  $\theta_d \sim Dirichlet(\alpha)$ <sup>2</sup>
  - For each word  $w_n \in d$  ( $N$  is the total number of words):
    - (a) Per-word topic assignment (shows which topic generated the word instance  $w_{d,n}$ ),  $z_{d,n} \sim Mult(\theta_d)$ <sup>3</sup>

---

<sup>1</sup>*Dirichlet*(.) is the Dirichlet distribution (a conjugate prior for the Multinomial distribution),  $\beta$  is a hyper-parameter

<sup>2</sup> $\alpha$  is a hyper-parameter.

<sup>3</sup>*Mult*(.) is the Multinomial distribution.

(b) An observed word,  $w_{d,n} \sim Mult(\varphi_k)$

The joint probability for LDA takes the form (2):

$$\begin{aligned} p(w_{d,n}, z_{d,n}, \theta_d, \varphi_k | \alpha, \beta) &= \left( \prod_{n=1}^N p(z_{d,n} | \theta_d) p(w_{d,n} | z_{d,n}, \varphi_{n,k}) \right) \left( \prod_{k=1}^K p(\varphi_k | \beta) \right) \left( \prod_{d=1}^D p(\theta_d | \alpha) \right) \\ &= \left( \prod_{n=1}^N Mult(z_{d,n} | \theta_d) Mult(w_{d,n} | z_{d,n}, \varphi_{d,k}) \right) \left( \prod_{k=1}^K Dirichlet(\varphi_k | \beta) \right) \left( \prod_{d=1}^D Dirichlet(\theta_d | \alpha) \right) \quad (2) \end{aligned}$$

where,  $p(w_{d,n}, z_{d,n}, \theta_d, \varphi_k | \alpha, \beta)$  is the posterior from the LDA model.

Latent variables  $z_{d,n}$ ,  $\theta_d$ ,  $\varphi_k$  are unobserved. Inference is done via Collapsed Gibbs Sampling (Griffiths & Steyvers (2004)) with  $\alpha = 50$  and  $\beta = 0.01$ . Since for the inference of both  $\theta_d$  and  $\varphi_k$  it is sufficient to know just  $z_{d,n}$ , Collapsed Gibbs Sampling is based on integrating out the multinomial parameters and simply sampling  $z_{d,n}$  (see Griffiths & Steyvers (2004) for the detailed treatment). The outcomes of the algorithm are topic distributions  $\theta_d$  and word distributions per topic  $\varphi_k$ .

The optimal number of topics for LDA was chosen based on coherence values. The topics are considered to be coherent if all or most of the words, for example, the topic's top N words, are related. Coherence values for different numbers of topics are presented in Figure A.1. According to the coherence values, the optimal number of topics is 40. All topics from the LDA model are interpretable and are shown in Figure 1, whereas Table A.1 shows word distributions for each topic.

Larsen & Thorsrud (2019) in their study implemented sign adjustment (positive versus negative news) to news topics. But, as was pointed out by Sims (2003), the tone of economic reporting affects sentiment beyond the economic information contained in reporting itself (which was explored in the study of Shapiro et al. (2017)). Therefore, I take into account both statement's topic and sentiments.

To assign a sentiment for each FOMC statement I employ a dictionary of Loughran & McDonald (2016) with a negation rule (details are discussed in Appendix B). This approach relates to Shapiro et al. (2017), where the authors found that a combination of different dictionaries with a negation rule is closer to human judgements in labelling sentiment.

Positive sentiment of a sentence is calculated as following (3):

$$Pos_i = \frac{\#positivewords_i - \#negativewords_i}{\#totalwords_i} \quad (3)$$

The total monthly positive sentiment for a certain economic topic is calculated as the



sum of sentence positive sentiments minus negative sentiments multiplied by topic proportions within a sentence and sum over the sentences (4):

Pos\_{topic} = \sum\_{i \in topic} Pos\_i \times topic\\_proportions\_i (4)

where topic\\_proportions\_i is the proportions of the topics in a sentence that is above a threshold (details can be found in Appendix B).



Figure 1: LDA topics

Similarly, I calculated uncertainty sentiments by employing (3) and (4) for uncertain

words from Loughran & McDonald (2016)<sup>4</sup>.

### 3 Information content of Fed communication

The pre-trained LDA model can be used to classify new documents. It decomposes any new document into forty topics by assigning topic proportions that sum up to one. Therefore, any document can be represented as forty topic proportions. These proportions should capture the meaning of a document. Appendix C presents the results for labelling topics for FOMC statements split by paragraphs and sentences. Topic distributions mainly correctly capture the meanings of each sentence and paragraph. Moreover, aggregated topic distributions over all documents are approximately the same in case of assigning a topic based on the threshold 0.3 for each sentence and 0.25 for each paragraph (see Figure C.19, Figure C.20 and Figure C.21).

Figure 2 shows aggregated topic distributions over all documents with topics assigned for each sentence. Based on the results, the Fed signals the most about its monetary policy (the Fed topic), economic conditions (Economic and Economics topics), federal committee regulations (the Rules topic), interest rates setting (the Rates topic), reporting (the Reports topic), job market conditions (the Jobs topic), asset market (the topics Investing and Securities), budget (the topic about income, taxes, budget and spendings), and oil/gas (the topic about gas, energy, oil prices, etc.).

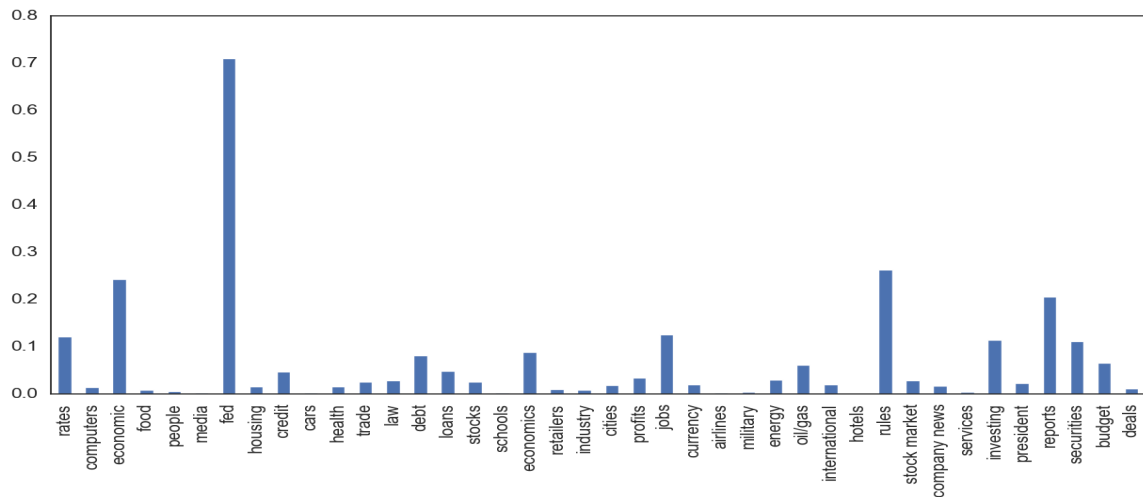


Figure 2: Topic proportions of statements by each sentence

<sup>4</sup>The full list of words for each sentiment category is available at <https://sraf.nd.edu/textual-analysis/resources/>

The topic decompositions for FOMC statements over time show that from 2008 the FOMC started to rely more on communications (Figure D.7). That is fully in line with that the federal funds rate hit the zero lower bound and the FOMC started to use unconventional monetary policy tools. The Fed started to communicate more about its monetary policy, but also about economic conditions, its interest rate settings, jobs, rules, reports, securities and investment.

The distribution of topics is also not constant over time. The FOMC releases more information about debt and loans from 2008, and about stocks and jobs from 2010. Additionally, the Fed communicates more about its interest rates policy from 2012 (Figure D.7). Moreover, the tone of the Fed during economic recessions is generally more negative (Appendix D).

I use the information contained in the FOMC statements to decompose monetary policy surprises into information and policy shocks. Surprises are changes in the federal funds futures on the dates of announcements in a narrow time window around these announcement<sup>5</sup>. To decompose surprises into information and non-information components I need to select the topics that are important for these surprises. Each FOMC statement is decomposed into 40 topics and not all information might be relevant for the public. I use a Bayesian Lasso regression (Park & Casella (2008)) for topic selections. For this purpose all non-stationary topic time series were transformed into a stationary form by taking first differences. All series were standardised for Lasso regression.

Figure 3 presents the Bayesian Lasso in the form of (1) for 40 topics time series from FOMC statements. It shows the proportions of samples when each topic was selected. The total number of MCMC samples is 10,000. One needs to set a threshold for selecting the most important topics. I use the threshold 0.65, so I select the topics that were included in at least 6,500 MCMC samples.

The topics that are found to be important for predicting Fed “surprises” are fully in line with what one would expect. These surprises are predicted from economic, credit, company news, inverting and deals topics. Moreover topics about monetary policy and international are important for “surprises” in 3-month federal funds futures. The results are also in line with the results of Jarocinski & Karadi (2020), who found that a difference between the staff and private forecasts about the one-quarter-ahead real GDP growth influences the

---

<sup>5</sup>Usually it is a 30-minutes window around the announcement time.

central bank information shocks significantly. The Lasso and Elastic net regression results confirm the findings from the Bayesian Lasso about important topics (Table E.1).

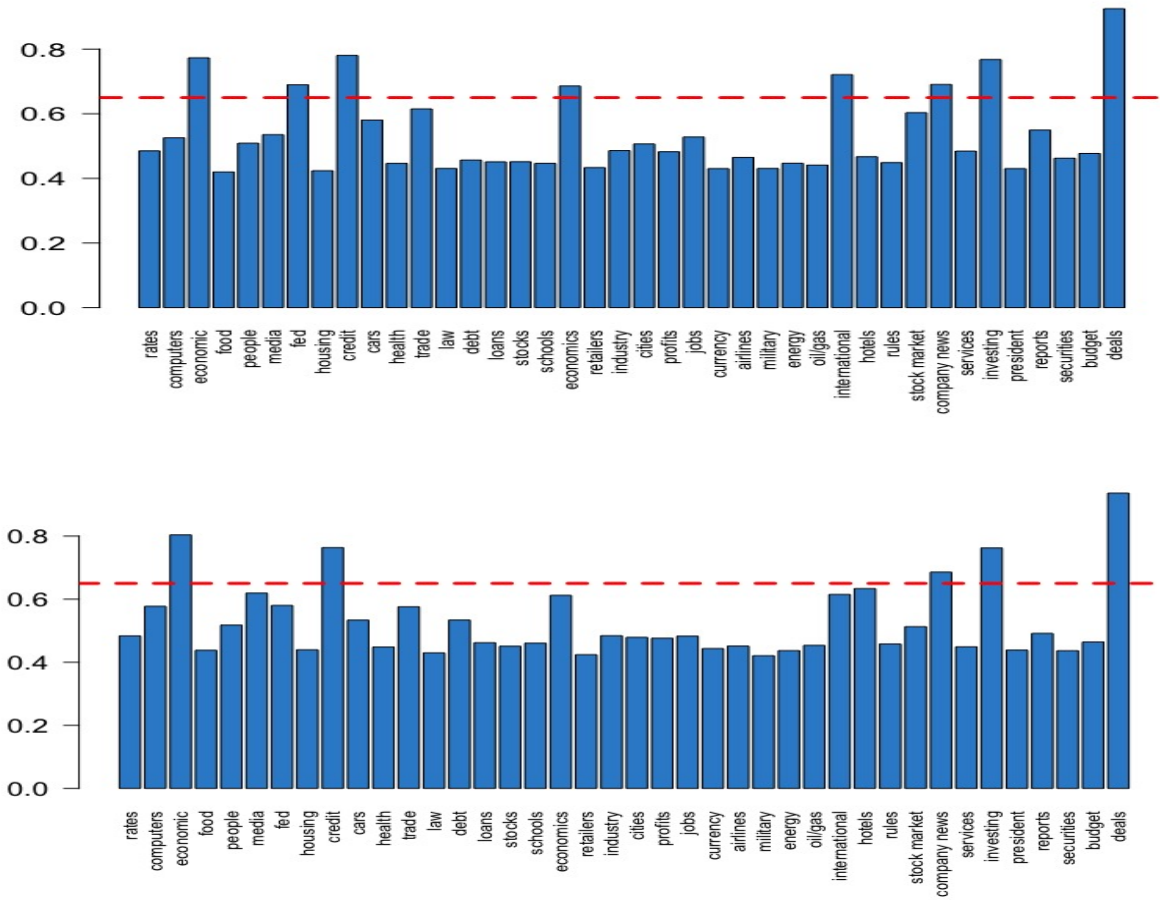


Figure 3: Bayesian Lasso for surprises in 3m federal funds futures (top) and the first principal component of the surprises in fed funds futures and eurodollar futures with one year or less to expiration (bottom)

Employing sign adjustment for topics from FOMC statements instead of tone adjustment leads to similar results, namely the topics Economic, Economics, Cities, Deals are important for surprises in federal funds futures on the FOMC statements release dates (Figure E.1). The Fed does not talk about cities in its statements. Rather the Cities is just a label for a topic from distribution of words. The topic Cities represents sentences that contain a certain combination of words, like citi, build, develop, offic, area, project, project, real, properti, million, estat, space, plan, squar, washington, district, construct, park, street,

local, leas, counti, feet, downtown, rent, land, region, commun, includ, commerci. It does not need to be about cities, but might be about development, projects and so on. This topic is quite infrequent in FOMC statements (Figure 2).

The topics about trade and industry with uncertainty sentiments are also found to be important for surprises in federal funds futures (Figure E.1).

On top of that, the topics Economic, Credit, Cars, Jobs, International, Company News, Investing and Deals are found to be important for Gertler & Karadi (2015) proxy for surprises in federal funds futures (Figure E.1). The topic Cars does not need to be about cars, but it is about car, sale, auto, vehicle, ford, year, motor, chrysler, truck, model, gm, gener, compani, dealer, market, product, automak, plant, industri, sold, sell, toyota, maker, unit, detroit, driver, incent, american, part, engin. This topic is also infrequent in FOMC statements.

## 4 Monetary policy vs. Information shocks

### 4.1 Baseline results

Following Jarocinski & Karadi (2020), I use Cholesky identification<sup>6</sup> for monetary policy shocks with Jarocinski & Karadi (2020) original variables in the following order: surprises in 3-month federal funds futures (alternatively, errors from the regression (1) as monetary policy shocks, and predicted values from (1) as information shocks), surprises in S&P500, the one-year government bond yield, real GDP, GDP deflator and the Excess Bond Premium. The studied period is 1994:M3–2016:M12. Appendix G presents the SVAR estimation details.

Figure 4 discusses the baseline results. Panel (a) presents the results where surprises in 3-month federal funds futures are ordered first, Panel (b) where errors from the regression (1) are ordered first, and Panel (c) where the predicted values from (1) are ordered first.

The baseline results (Panel (a)) are in line with the results of Jarocinski & Karadi (2020) except for S&P500 responses. This might be explained by different studied period since Jarocinski & Karadi (2020) used the period from 1984 and employed Kalman filter and smoother for filling missing values in surprises in 3-month federal funds futures. The

---

<sup>6</sup>The authors use Cholesky identification as alternative specifications to sign restrictions.

responses of the one year rate and the Excess Bond Premium have smaller magnitudes than in Jarocinski & Karadi (2020).

Panel (b) presents the results for purged shocks, that should be free of the Fed information effect. The results are similar to Jarocinski & Karadi (2020), except that the response of EBP has the same magnitude as in Panel (a). The response of the one year rate is more transitory than in Panel (a). The response of S&P500 is negative for the first few months. The response of real GDP has greater magnitude and it is more prolonged. Finally, the response of GDP deflator is more prolonged compared to the results in Panel (a). Therefore, the identified effect looks like a contractionary monetary policy shock.

Panel (c) discusses the results for information shocks. The results are in line with Jarocinski & Karadi (2020) and Steinsson (2019): the Fed information shock has a more prolonged effect on the one year rate, positive effect on S&P500, positive long-run effect on real GDP and negative effect on the EBP. The main difference to the results of Jarocinski & Karadi (2020) lies in a more positive long-run response of real GDP. Moreover, the response of surprises in S&P500 is negative, instead of being positive, but has smaller magnitude than in response to a policy shock.

The result of a smaller decline in S&P500 surprise in a tight window can be explained in line with Steinsson (2019), who stated that a pure tightening of monetary policy leads stock prices to fall for two reasons: higher discount rates and lower output. The authors found that if monetary policy conveys information about both future monetary policy and future exogenous economic fundamentals, stock prices fall by less amount in response to the FOMC announcement than to the shock without information about future exogenous fundamentals.

The differences to Jarocinski & Karadi (2020) might be explained by (1) different periods studied<sup>7</sup>, (2) different identification strategies for monetary policy and information shocks<sup>8</sup>.

---

<sup>7</sup>Jarocinski & Karadi (2020) dealt with missing values for the shocks series via Kalman filter and smoother.

<sup>8</sup>Jarocinski & Karadi (2020) employed sign restriction which is set identification while Cholesky is point identification.

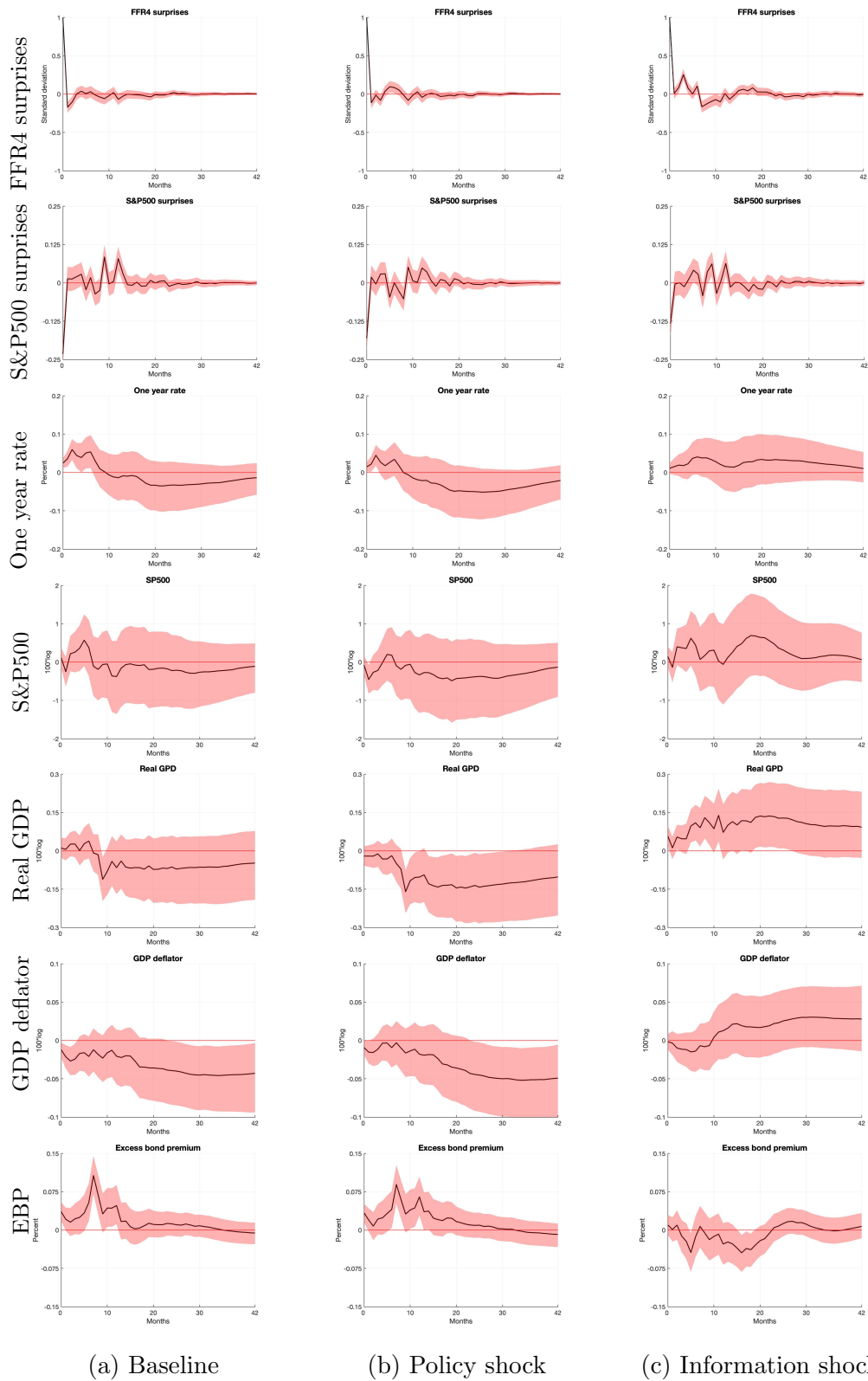


Figure 4: Comparison between monetary policy and information shocks. 3m federal funds futures shaded 16% and 84% percentiles

Appendix H discusses the results using topics from FOMC statements with sign adjustment instead of tone adjustment. I use the topics with sign adjustment that were found to be important for surprises in 3-month federal funds futures (Figure E.1 Panel (a)) and employ (1) to purge surprises in federal funds futures with respect to sign adjusted topics. The results are similar to the presented above except for a more muted response of real GDP and inflation to a policy shock (Panel (b)). The responses of the one year rate and inflation to an information shock have larger magnitudes (Panel (c)).

## 4.2 Robustness analysis

For the robustness check I use the first principal component of surprises in the current month and 3-month fed funds futures and 2-, 3-, and 4- quarters ahead 3-month eurodollar futures (Jarocinski & Karadi (2020)). I purge this series in a similar way to the previous one but use topics with tone adjustment that were found to be important for this principal component (Figure 3 bottom panel). I also use a measure of the stock price surprises from Jarocinski & Karadi (2020), which is the first principal component of the surprises in the S&P500, Nasdaq Composite and Wilshire 5000. Figure 5 presents the results.

The results are similar to the results from Figure 4, except for a more muted response in inflation. The effects of a policy shock on real GDP and one year rate are completely in line with the previous findings.

Another difference with previous findings lies in a larger effect of information shock on the one year rate. The magnitude of the effect of information shock is also larger for S&P500 and real GDP compared to the findings using surprises in 3-month federal funds futures.



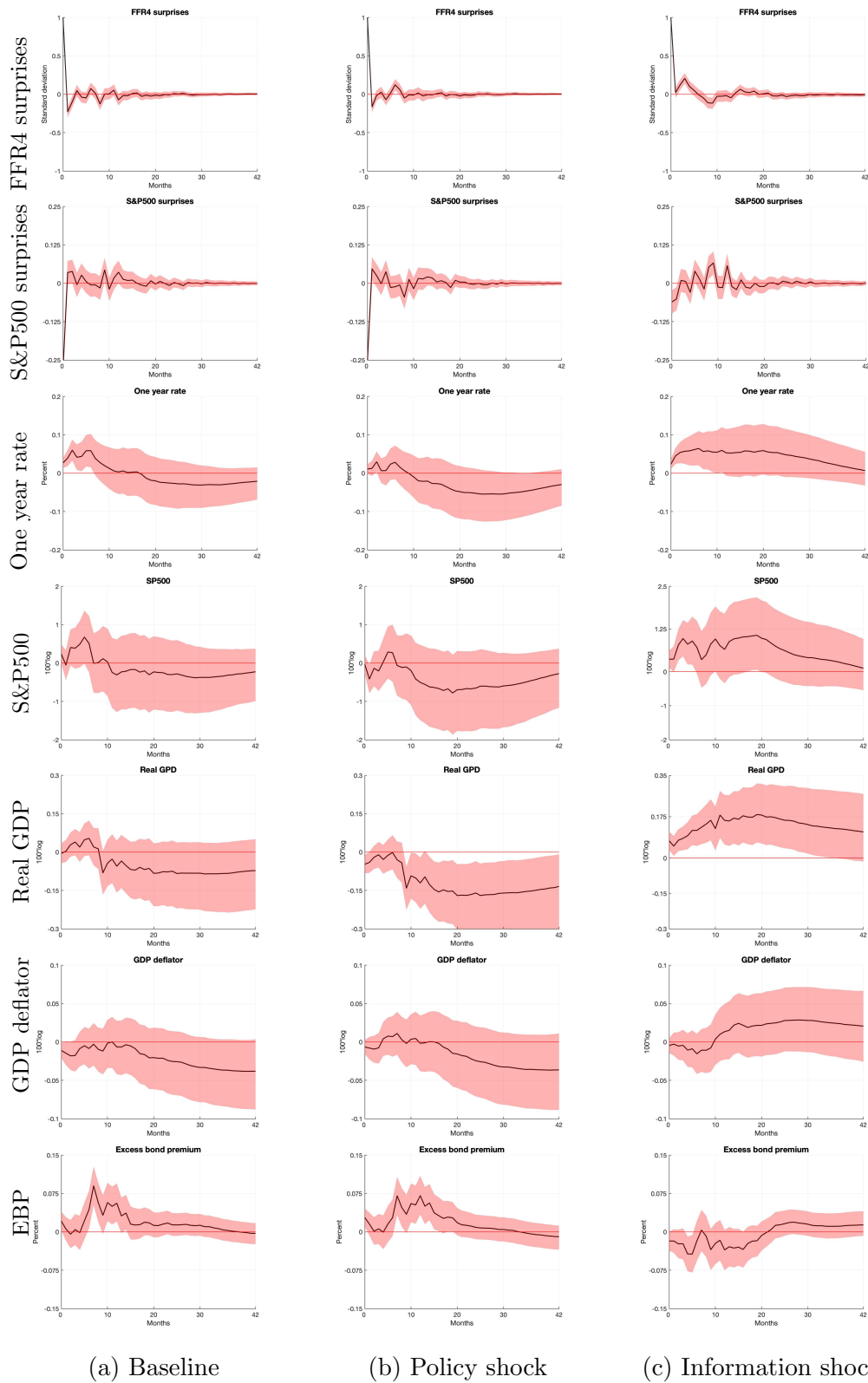


Figure 5: Comparison between monetary policy and information shocks. The first principal component of the surprises in fed funds futures and eurodollar futures with one year or less to expiration shaded 16% and 84% percentiles

## 5 Conclusions

The paper elaborates on the recent contribution of Jarocinski & Karadi (2020) in decomposing information from policy shocks. This study uses the information from FOMC statements and directly decomposes surprises in 3-month federal funds futures into a part that is explained by this information and a part that is not. I extract information from FOMC statements by using Latent Dirichlet Allocation that was pre-trained on the business section from major US newspapers.

The study combines topic time series from FOMC statements with the tone of these statements. This tone was assigned using a lexicon-based approach that counts positive and negative words in each sentence. After topics time series were adjusted for the tone, these series were investigated by their predictive power for surprises in 3-month federal funds futures on the FOMC meeting dates. The topics, that were found to be important for these surprises, are about the economy, credit, investment, company news and deals.

After decomposing surprises in Fed futures into a part that is explained by topics from Fed statements and that is not explained, the study employs these purged series as proxies for monetary policy and Fed information shocks. The results show that a policy shock has a more negative effect on GDP and a more prolonged negative effect on inflation compared to the baseline surprises measure. In the short-run it causes S&P500 to decline and the Fed to raise its interest rate. A Fed information shock has positive long-run effects on S&P500, on the interest rate, on real GDP, and a negative short-run effect on inflation. Moreover, it reduces the costs of credit.

## References

- Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003), ‘Latent Dirichlet Allocation’, *Journal of Machine Learning Research* **3**, 993–1022.
- Darling, W. M. (2011), ‘A Theoretical and Practical Implementation Tutorial on Topic Modeling and Gibbs Sampling’, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* pp. 642–647.
- Federal Open Market Committee: Transcripts and other historical materials* (2019).  
**URL:** [https://www.federalreserve.gov/monetarypolicy/fomc\\_historical.htm](https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm)
- Federal Reserve Economic Data* (2019).  
**URL:** <https://fred.stlouisfed.org>
- Gertler, M. & Karadi, P. (2015), ‘Monetary Policy Surprises, Credit Costs, and Economic Activity’, *American Economic Journal: Macroeconomics* **7**(1), 44–76.
- Greenbook Historical and Forecast Data* (2019).  
**URL:** <https://www.philadelphiafed.org/research-and-data/real-time-center/greenbook-data/philadelphia-data-set>
- Griffiths, T. & Steyvers, M. (2004), ‘Finding scientific topics’, *Proceedings of the National Academy of Sciences* **101**(1), 5228–5235.
- Hansen, S. & McMahon, M. (2016), ‘Shocking language: understanding the macroeconomic effects of central bank communication’, *CAMA Working Paper* (4/2016).
- Hansen, S., McMahon, M. & Tong, M. (2019), ‘The long-run information effect of central bank communication’, *Journal of Monetary Economics* **108**, 185–202.
- Hubert, P. (2019), ‘State-Dependent Effects of Monetary Policy: the Central Bank Information Channel’, *Sciences Po OFCE Working Paper* (04/2019).
- Hubert, P. & Labondance, F. (2017), ‘Central bank sentiment and policy expectations’, *Bank of England Staff Working Paper* (648).

- Iglesias, J., Ortiz, A. & Rodrigo, T. (2017), ‘How do the Central Banks talk?: A Big Data approach to Turkey’, *BBVA Bank, Economic Research Department Working Papers* (17/24).
- Jarocinski, M. & Karadi, P. (2020), ‘Deconstructing monetary policy surprises: the role of information shocks’, *American Economic Journal: Macroeconomics* **12**(2), 1–43.
- Koop, G. & Korobilis, D. (2010), ‘Bayesian Multivariate Time Series Methods for Empirical Macroeconomics’, *Foundations and Trends in Econometrics* **3**(4), 267–358.
- Larsen, V. & Thorsrud, L. (2019), ‘The Value of News for economic developments’, *Journal of Econometrics* **210**(1), 203–218.
- Loughran, T. & McDonald, B. (2016), ‘Textual Analysis in Accounting and Finance: A Survey’, *Journal of Accounting Research* **54**(4), 1187–1230.
- Miranda-Agrippino, S. & Ricco, G. (2014), ‘The Transmission of Monetary Policy shocks’, *Bank of England Staff Working Paper* (657).
- Park, T. & Casella, G. (2008), ‘The Bayesian Lasso’, *Journal of the American Statistical Association* **103**(482), 681–686.
- Romer, C. D. & Romer, D. H. (2000), ‘Federal Reserve Information and the Behavior of Interest Rates’, *The American Economic Review* **90**(3), 429–457.
- Romer, C. D. & Romer, D. H. (2004), ‘A New Measure of Monetary Shocks: Derivation and Implications’, *The American Economic Review* **94**(4), 1055–1084.
- Shapiro, A. H., Sudhof, M. & Wilson, D. (2017), ‘Measuring News Sentiment’, *Federal Reserve Bank of San Francisco Working Paper Series Working Paper* (2017-01).
- Sims, C. A. (2003), ‘Implications of rational inattention’, *Journal of Monetary Economics* **50**(3), 665–690.
- Steinsson, E. N. J. (2019), ‘High Frequency Identification of Money Non-Neutrality: the Information effect’, *NBER Working Paper* (19260).

## Appendix A. Latent Dirichlet Allocation

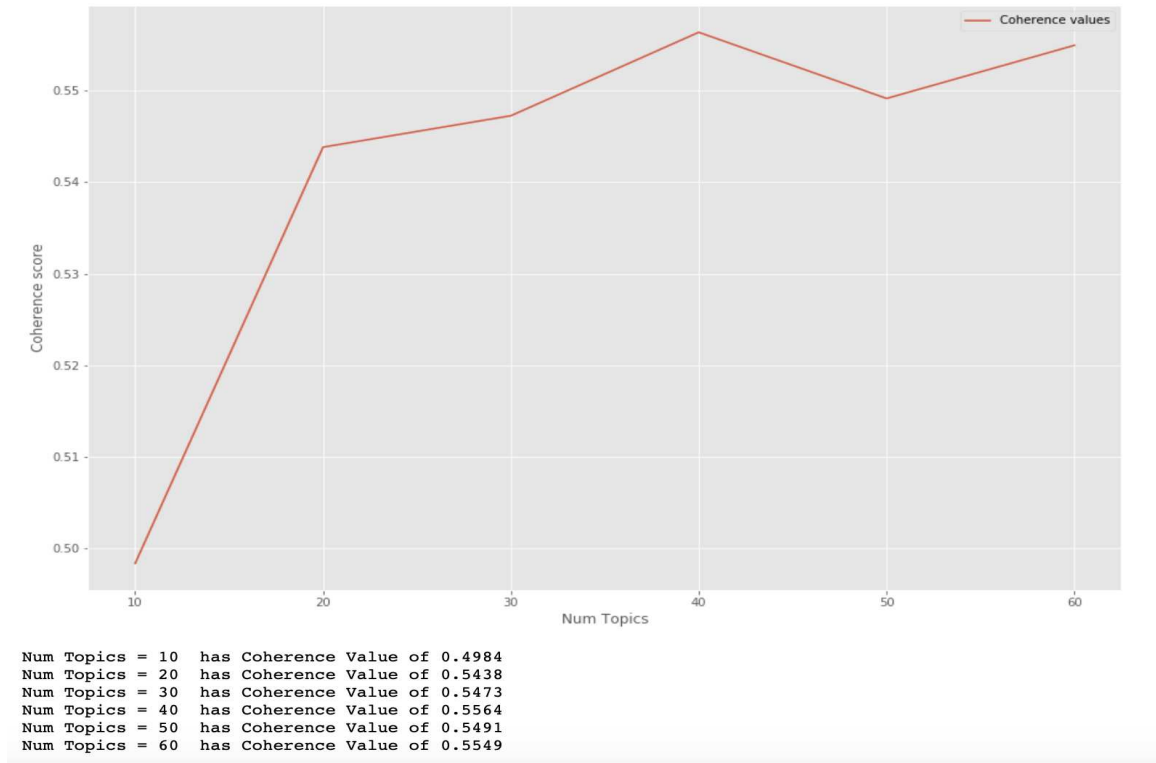


Figure A.1: Coherence values for the number of topics

Table A.1: Topic labelling for the LDA model

Topic	Words
rates	percent, year, increas, rate, averag, price, declin, rise, month, drop
computers	comput, technolog, compani, system, softwar, product, appl, microsoft, electron, market
economic	year, economi, growth, market, recess, expect, econom, mani, continu, industri
food	food, year, product, price, farm, market, farmer, restaur, agricultur, produc
people	peopl, time, make, thing, day, good, lot, work, back, tri
media	advertis, onlin, ad, site, internet, web, time, media, googl, publish
fed	rate, fed, interest, inflat, feder, reserv, economi, econom, polici, economist
housing	home, hous, california, lo, angel, year, price, counti, sale, san
credit	credit, consum, card, pay, custom, fee, account, servic, charg, check
cars	car, sale, auto, vehicl, ford, year, motor, chrysler, truck, model
health	insur, health, drug, care, compani, cost, medic, hospit, plan, year
trade	trade, state, unit, american, countri, foreign, import, world, mexico, export
law	case, court, investig, file, law, feder, charg, lawyer, attorney, judg
debt	debt, financi, billion, govern, bankruptci, crisi, plan, financ, money, problem
loans	bank, loan, mortgag, financi, feder, save, institut, borrow, lender, lend
stocks	stock, market, index, point, dow, rose, fell, gain, close, share
schools	chicago, school, photo, student, illinoi, famili, univers, colleg, program, tribun
economics	studi, econom, research, chang, univers, professor, differ, mani, exampl, problem
retailers	store, retail, sale, shop, year, chain, custom, buy, consum, holiday
industry	compani, industri, product, manufactur, steel, million, busi, produc, equip, oper
cities	citi, build, develop, offic, area, project, project, real, properti, million
profits	million, quarter, share, billion, earn, year, profit, compani, cent, sale
jobs	job, worker, work, employ, labor, employe, union, wage, unemploy, peopl
currency	dollar, york, cent, price, gold, trade, late, exchang, futur, currenc
airlines	airlin, travel, unit, air, fare, american, flight, carrier, boe, airport
military	war, govern, nation, countri, offici, attack, militari, soviet, world, defens
energy	power, energi, electr, state, util, plant, ga, water, cost, project
oil/gas	price, oil, energi, barrel, ga, product, gasolin, crude, day, produc
international	global, european, world, unit, europ, china, countri, british, intern, bank
hotels	hotel, photo, room, year, park, show, game, open, peopl, time
rules	propos, rule, regul, agenc, offici, feder, requir, law, member, committe
stock market	trade, market, stock, exchang, firm, secur, street, wall, futur, option
company news	compani, busi, execut, chief, firm, manag, presid, corpor, offic, year
services	servic, compani, comun, phone, network, custom, provid, busi, cabl, telephon
investing	fund, invest, stock, investor, market, manag, money, return, year, valu
president	presid, hous, republican, democrat, obama, trump, senat, white, polit, administr
reports	report, month, consum, economist, depart, increas, rose, declin, good, show
securities	bond, rate, treasuri, market, yield, price, issu, interest, note, secur
budget	tax, incom, year, budget, cut, plan, spend, save, pay, benefit
deals	compani, share, deal, million, offer, stock, billion, sharehold, merger, bid

## Appendix B. Tone adjustment for topic time series

For assigning a sentiment for each sentence from FOMC statements I use a negation rule. If the following words precede a collocation in the three-word window, then they are labelled as an opposite sentiment. Negation dictionary consists of the following words: aint, arent, cannot, cant, couldnt, darent, didnt, doesnt, ain't, aren't, can't, couldn't, daren't, didn't, doesn't, dont, hadnt, hasnt, havent, isnt, mightnt, mustnt, neither, don't, hadn't, hasn't, haven't, isn't, mightn't, mustn't, neednt, needn't, never, none, nope, nor, not, nothing, nowhere, oughtnt, shant, shouldnt, wasnt, werent, oughtn't, shan't, shouldn't, wasn't, weren't, without, wont, wouldnt, won't, wouldn't, rarely, seldom, despite, no, nobody.

I assign tone for each sentence based on three different strategies:

1. Positivity is calculated for each sentence and it scales its topic frequencies which are higher than the threshold (0.3).
2. Sign (positive/negative) is calculated for each sentence and it scales its topic frequencies which are higher than the threshold (0.3).
3. Uncertainty is calculated for each sentence and it scales its topic frequencies which are higher than the threshold (0.3).

## Appendix C. LDA and Fed Statements

### C.1 Performance of LDA by paragraphs

1. The federal reserve board today approved an increase in the discount rate from 4 3/4 percent to 5 1/4 percent, effective immediately. 1995-02-01

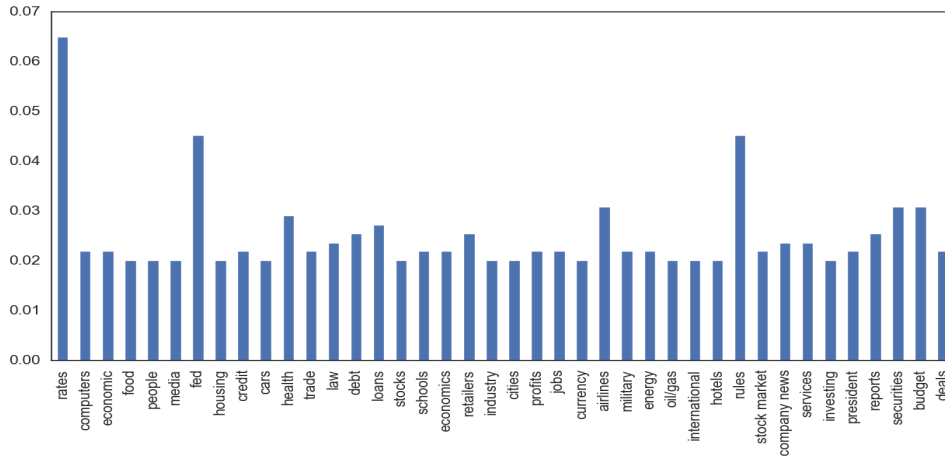


Figure C.1: Topic proportions for the paragraph 1

2. The committee perceives the upside and downside risks to the attainment of sustainable growth for the next few quarters are roughly equal. the probability of an unwelcome fall in inflation has diminished in recent months and now appears almost equal to that of a rise in inflation. with inflation quite low and resource use slack, the committee believes that it can be patient in removing its policy accommodation. 2004-03-16

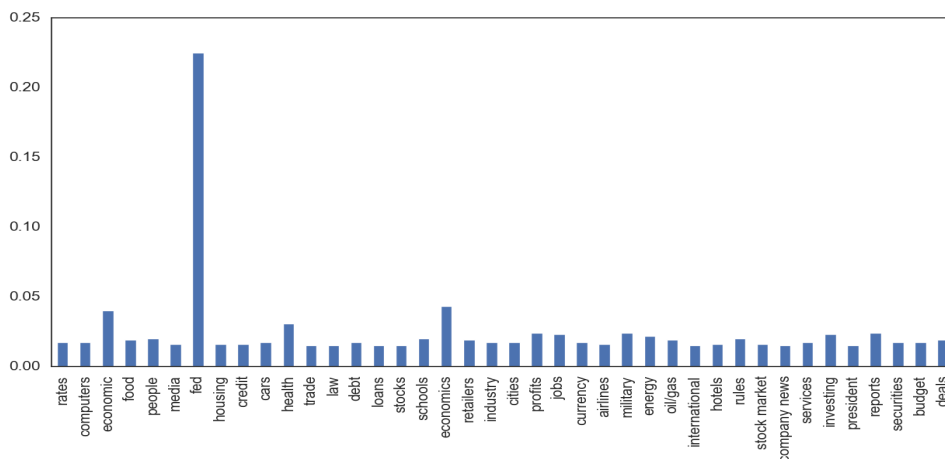


Figure C.2: Topic proportions for the paragraph 2



3. Developments in financial markets since the committee’s last regular meeting have increased the uncertainty surrounding the economic outlook. the committee will continue to assess the effects of these and other developments on economic prospects and will act as needed to foster price stability and sustainable economic growth. 2007-09-18

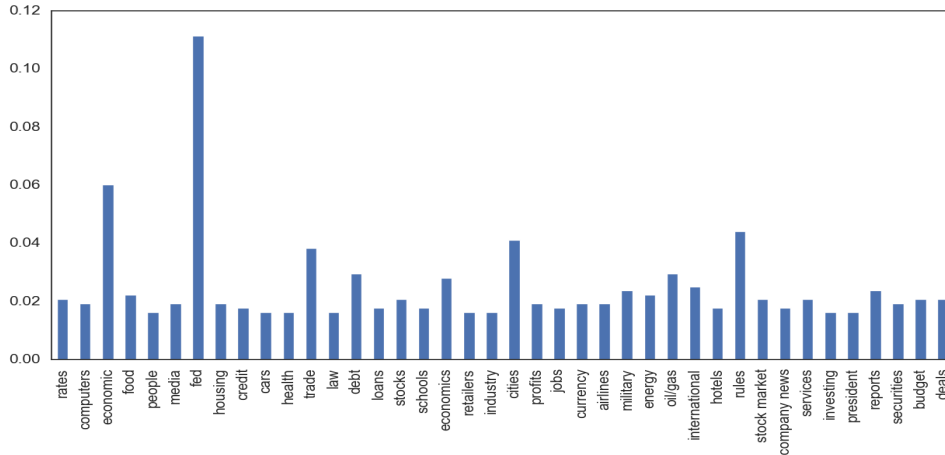


Figure C.3: Topic proportions for the paragraph 3

4. Strains in financial markets have increased significantly and labor markets have weakened further. economic growth appears to have slowed recently, partly reflecting a softening of household spending. tight credit conditions, the ongoing housing contraction, and some slowing in export growth are likely to weigh on economic growth over the next few quarters. over time, the substantial easing of monetary policy, combined with ongoing measures to foster market liquidity, should help to promote moderate economic growth. 2008-09-16

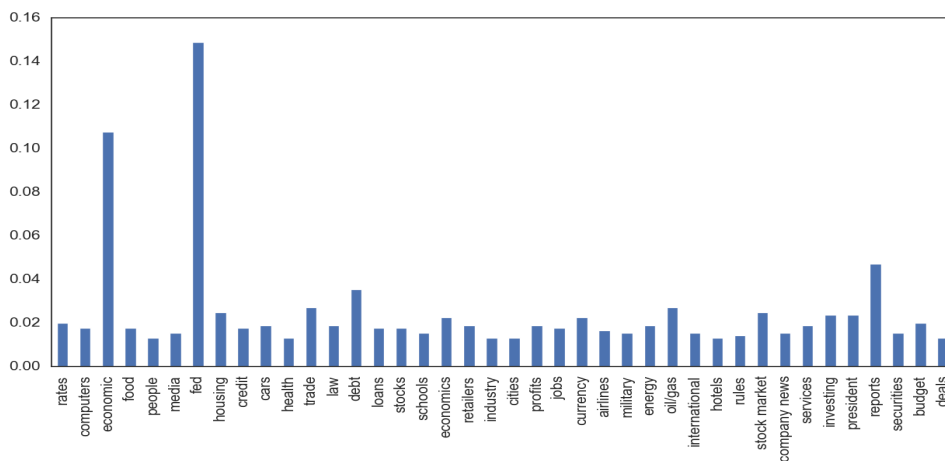


Figure C.4: Topic proportions for the paragraph 4

5. Inflation has been high, spurred by the earlier increases in the prices of energy and some other commodities. the committee expects inflation to moderate later this year and next year, but the inflation outlook remains highly uncertain. 2008-09-16

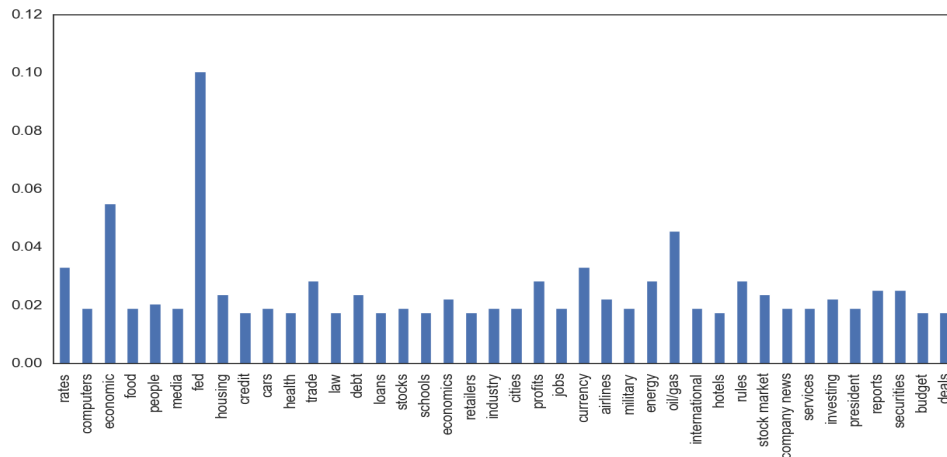


Figure C.5: Topic proportions for the paragraph 5

6. The downside risks to growth and the upside risks to inflation are both of significant concern to the committee. the committee will monitor economic and financial developments carefully and will act as needed to promote sustainable economic growth and price stability. 2008-09-16

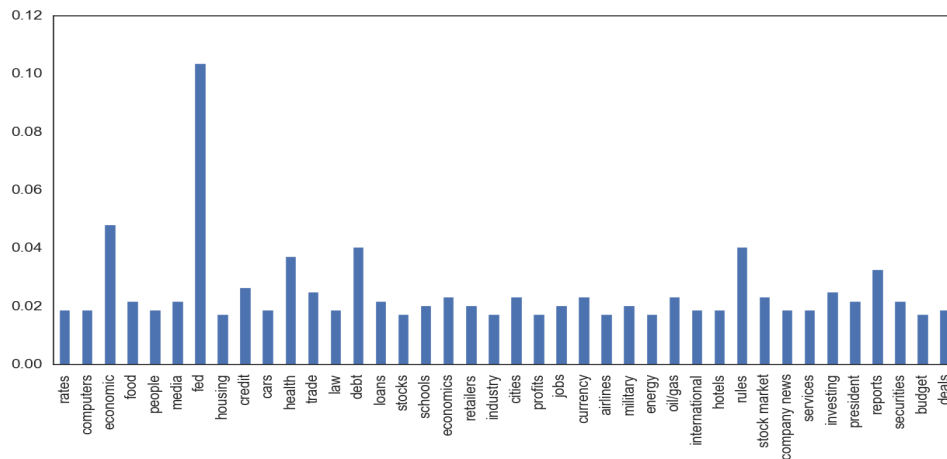


Figure C.6: Topic proportions for the paragraph 6

7. Throughout the current financial crisis, central banks have engaged in continuous close consultation and have cooperated in unprecedented joint actions such as the provision of liquidity to reduce strains in financial markets. 2008-10-08

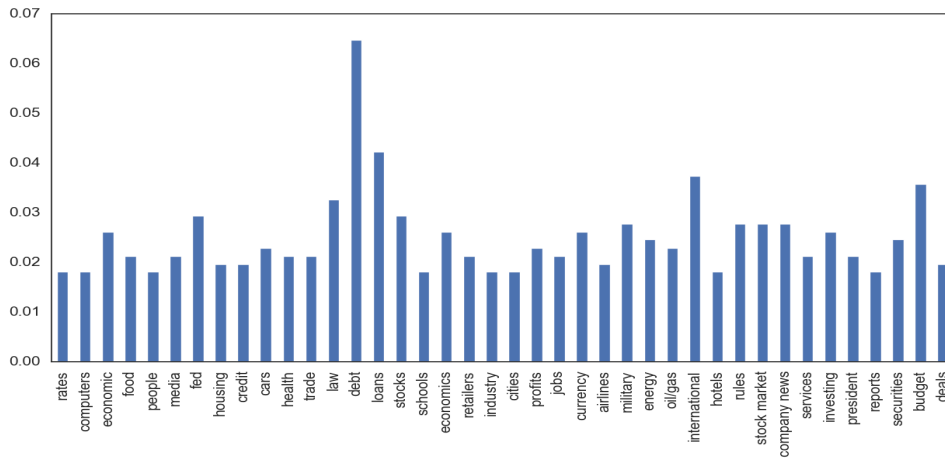


Figure C.7: Topic proportions for the paragraph 7

8. Information received since the federal open market committee met in june indicates that the labor market strengthened and that economic activity has been expanding at a moderate rate. job gains were strong in june following weak growth in may. on balance, payrolls and other labor market indicators point to some increase in labor utilization in recent months. household spending has been growing strongly but business fixed investment has been soft. inflation has continued to run below the committee’s 2 percent longer-run objective, partly reflecting earlier declines in energy prices and in prices of non-energy imports. market-based measures of inflation compensation remain low; most survey-based measures of longer-term inflation expectations are little changed, on balance, in recent months. 2016-07-27

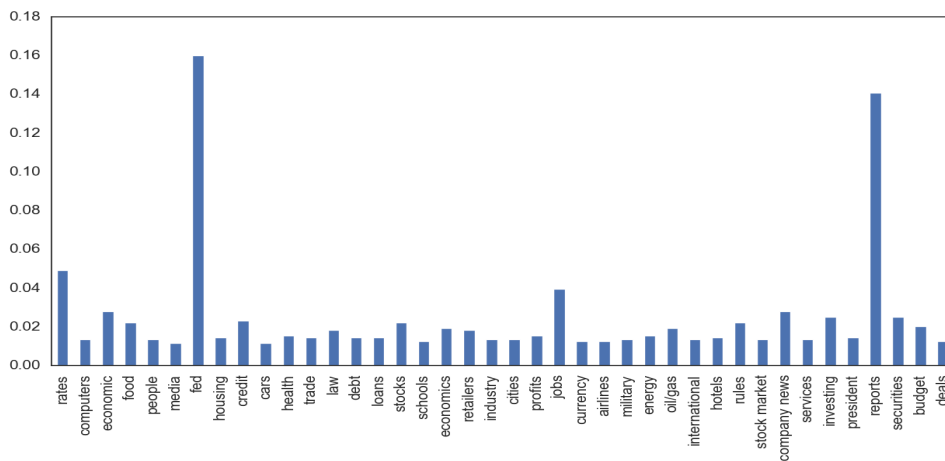


Figure C.8: Topic proportions for the paragraph 8

9. The coronavirus outbreak is causing tremendous human and economic hardship across the united states and around the world. the virus and the measures taken to protect public health are inducing sharp declines in economic activity and a surge in job losses. weaker demand and significantly lower oil prices are holding down consumer price inflation. the disruptions to economic activity here and abroad have significantly affected financial conditions and have impaired the flow of credit to u.s. households and businesses. 2020-04-29

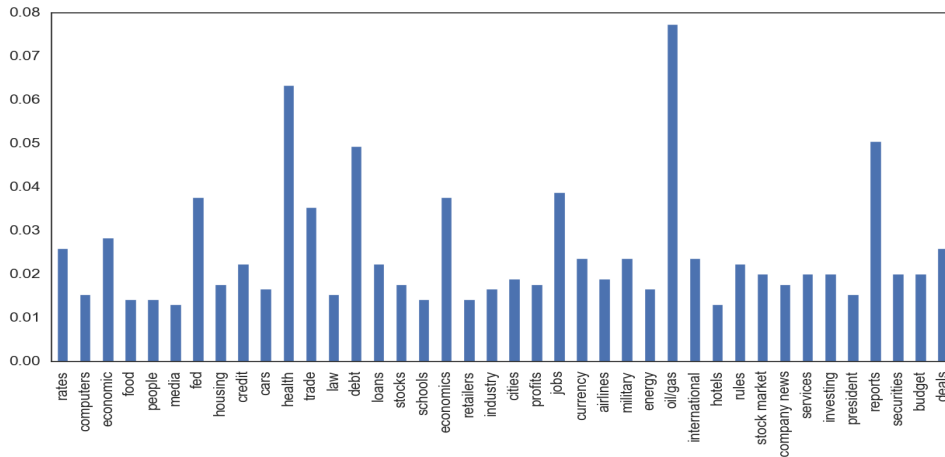


Figure C.9: Topic proportions for the paragraph 9

## C.2 Performance of LDA by sentences

1. Job gains have been strong, on average, in recent months, and the unemployment rate has remained low. 2018-12-19

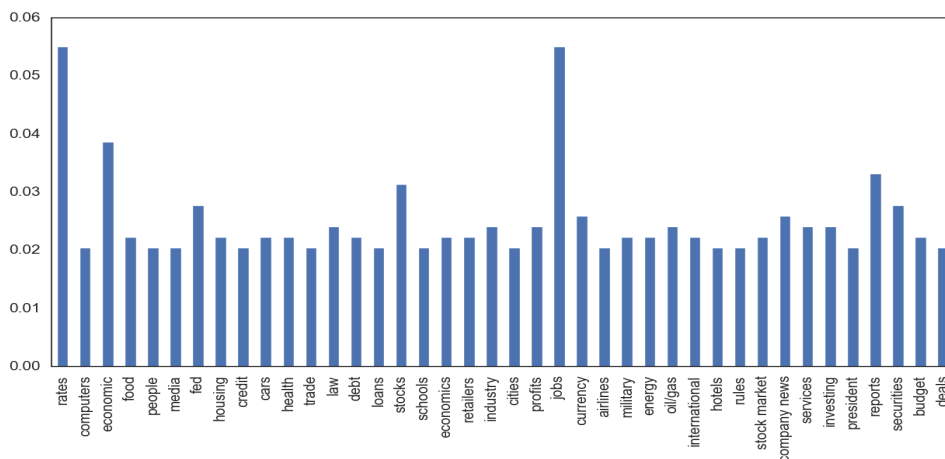


Figure C.10: Topic proportions for the sentence 1

2. Household spending has continued to grow strongly, while growth of business fixed investment has moderated from its rapid pace earlier in the year. 2018-12-19

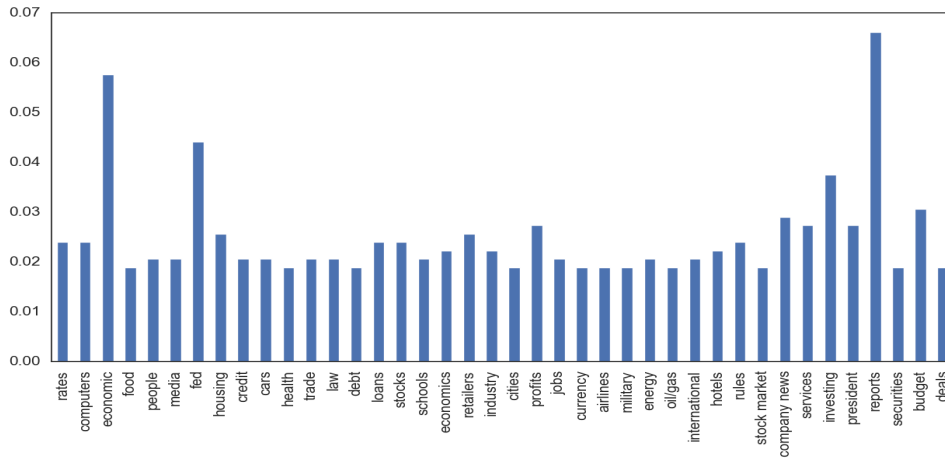


Figure C.11: Topic proportions for the sentence 2

3. On a 12-month basis, both overall inflation and inflation for items other than food and energy remain near 2 percent. 2018-12-19

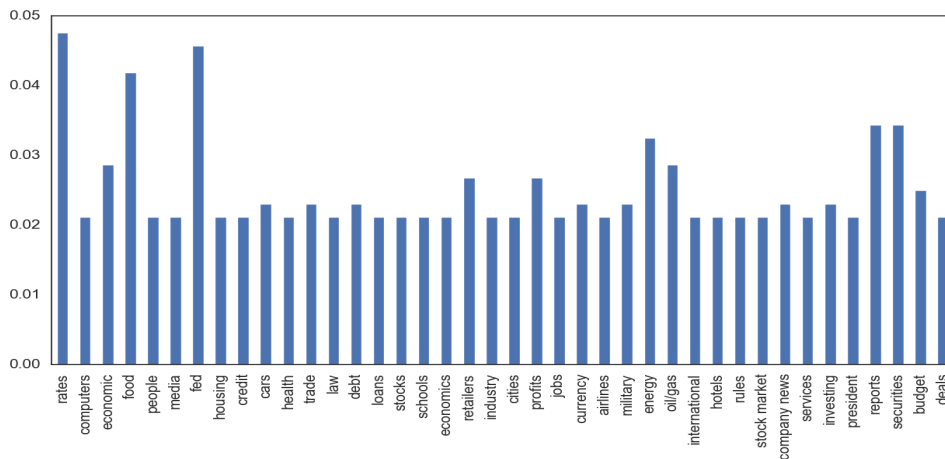


Figure C.12: Topic proportions for the sentence 3

4. Indicators of longer-term inflation expectations are little changed, on balance. 2018-12-19

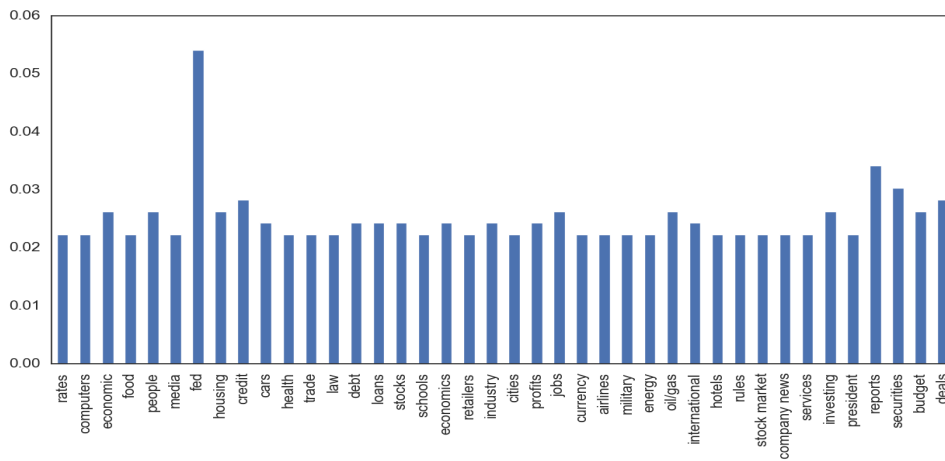


Figure C.13: Topic proportions for the sentence 4

5. Consistent with its statutory mandate, the committee seeks to foster maximum employment and price stability. 2018-12-19

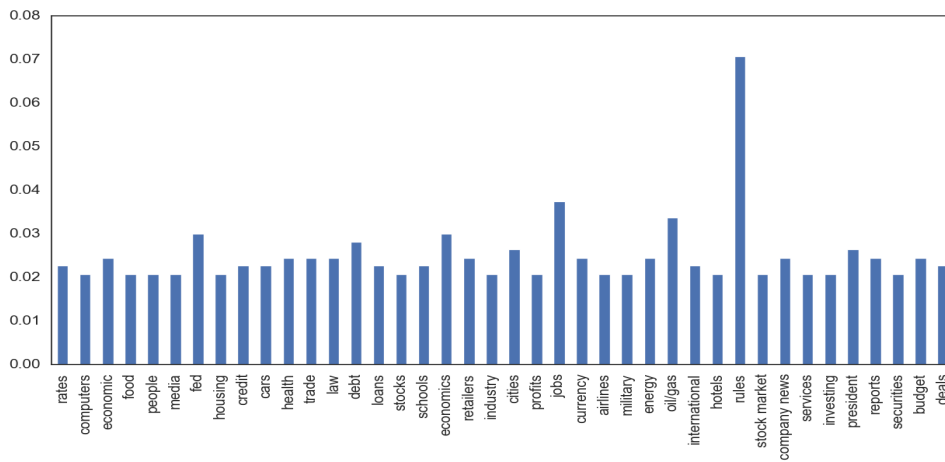


Figure C.14: Topic proportions for the sentence 5

6. The committee judges that some further gradual increases in the target range for the federal funds rate will be consistent with sustained expansion of economic activity, strong labor market conditions, and inflation near the committee's symmetric 2 percent objective over the medium term. 2018-12-19

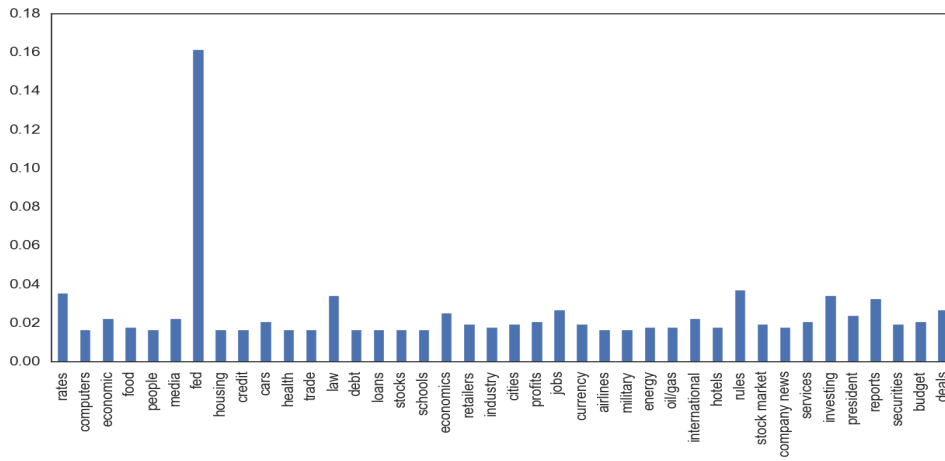


Figure C.15: Topic proportions for the sentence 6

7. The committee judges that risks to the economic outlook are roughly balanced, but will continue to monitor global economic and financial developments and assess their implications for the economic outlook. 2018-12-19

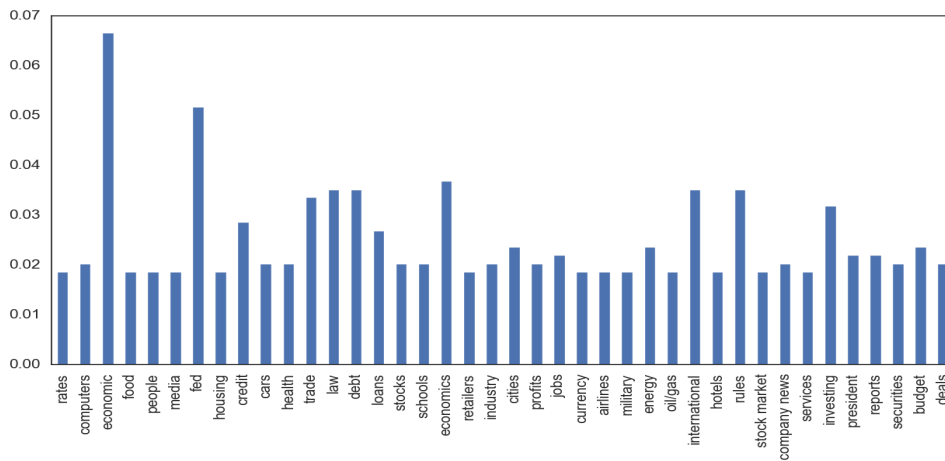


Figure C.16: Topic proportions for the sentence 7

8. In view of realized and expected labor market conditions and inflation, the committee decided to raise the target range for the federal funds rate to 2-1/4 to 2-1/2 percent. 2018-12-19

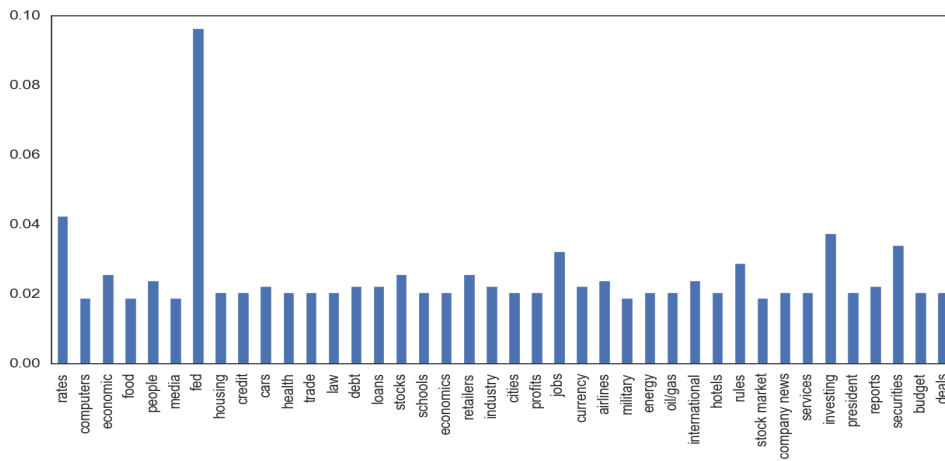


Figure C.17: Topic proportions for the sentence 8

9. 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. 2018-12-19

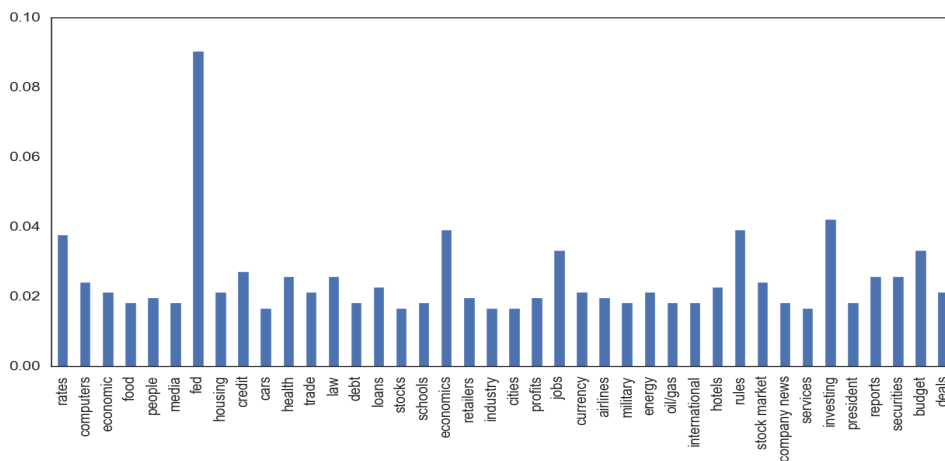


Figure C.18: Topic proportions for the sentence 9



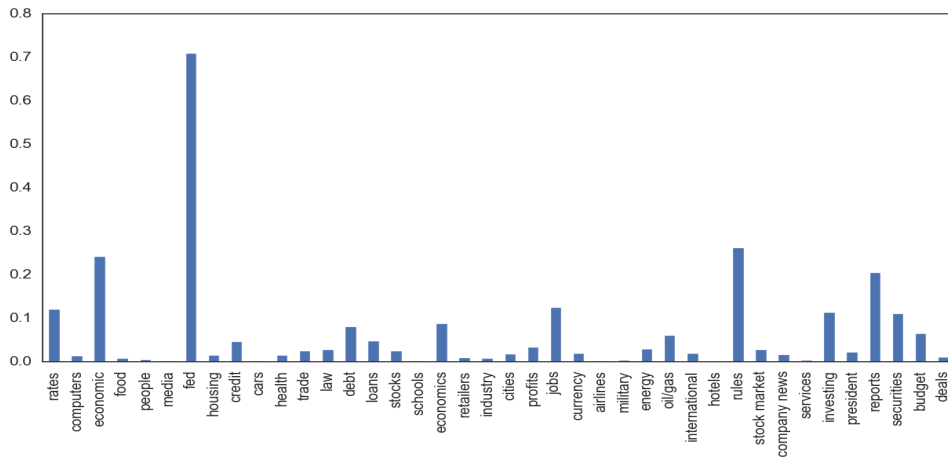


Figure C.19: Aggregated topic proportions by sentence

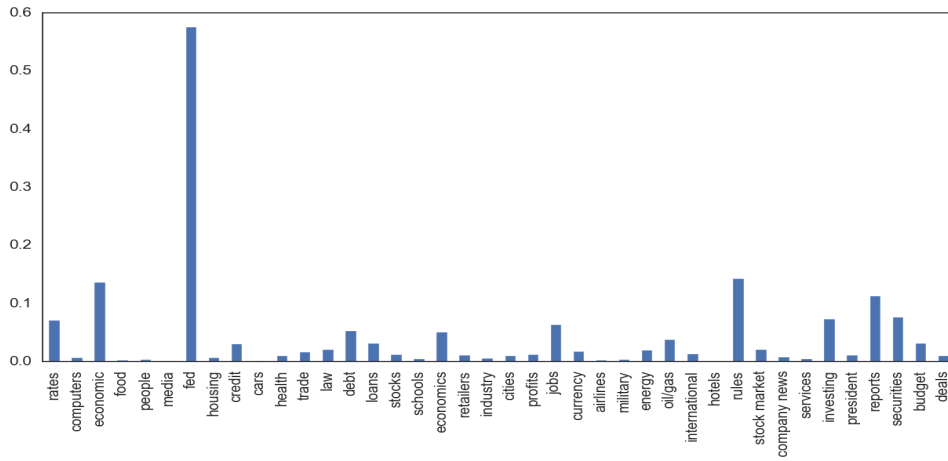


Figure C.20: Aggregated topic proportions by paragraph

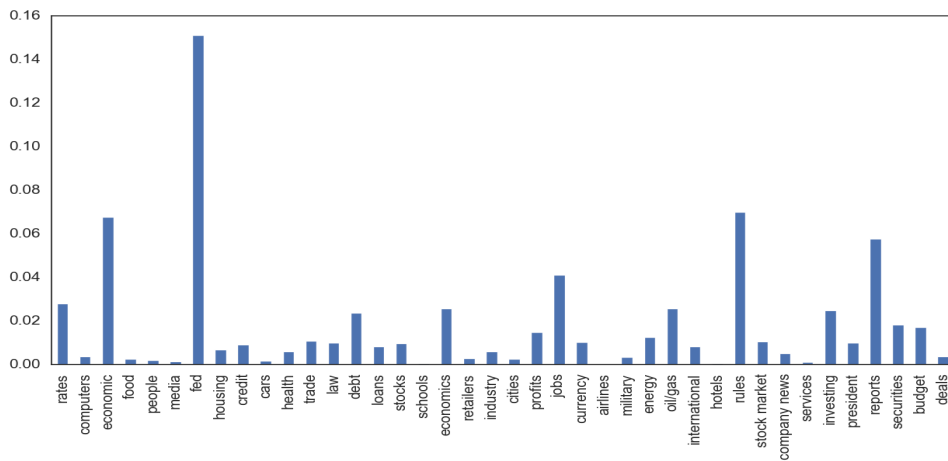


Figure C.21: Aggregated topic proportions by sentence with sign adjustment

## Appendix D. Information in FOMC statements

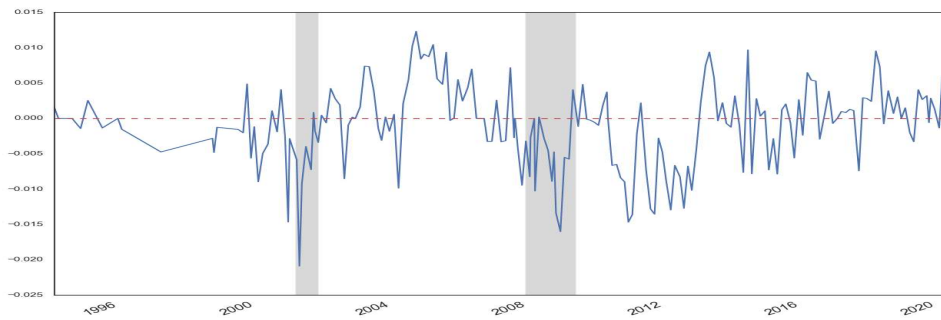


Figure D.1: Economic topic

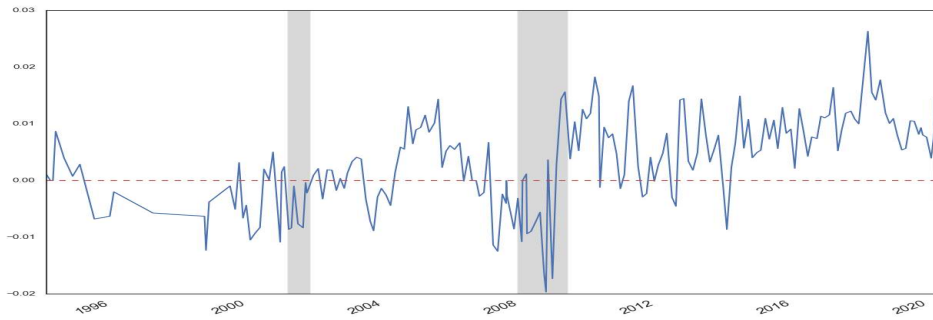


Figure D.2: Fed topic

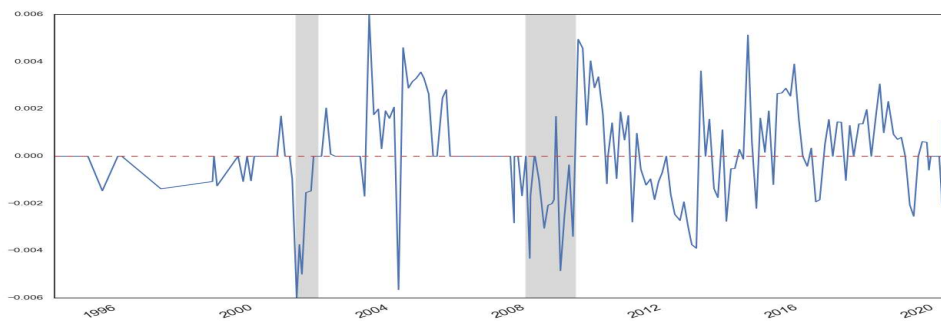


Figure D.3: Investment topic

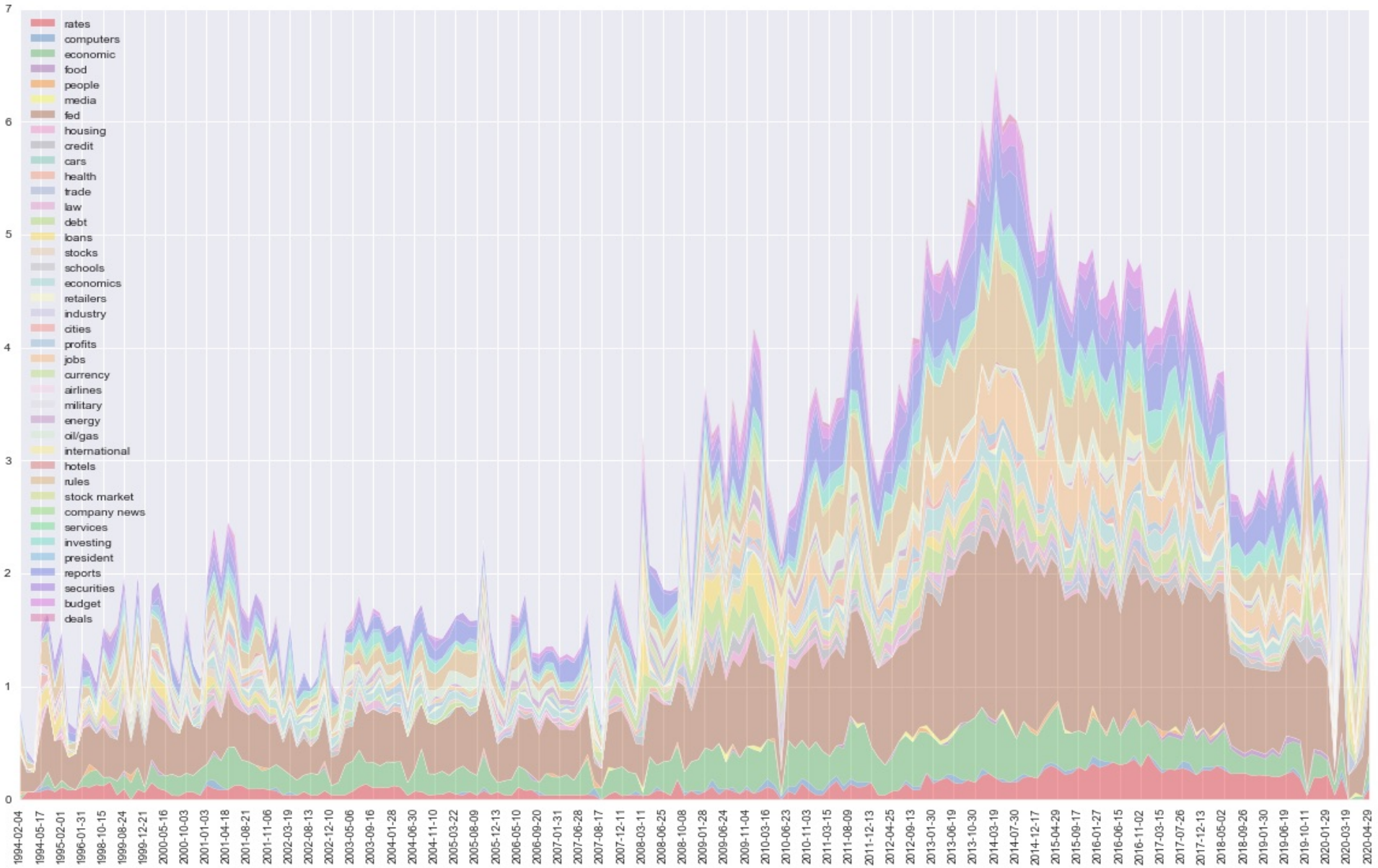
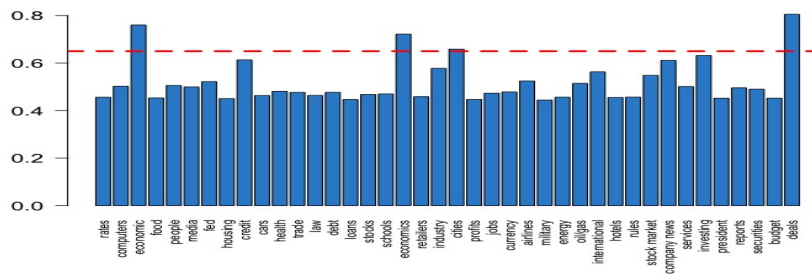
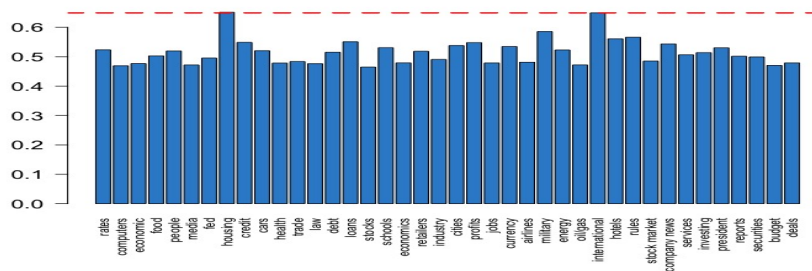


Figure D.4: Topic frequencies over time

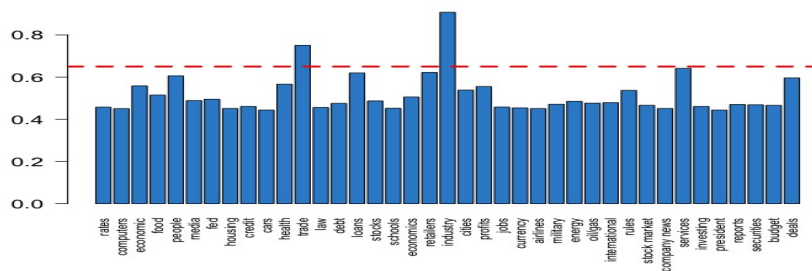
## Appendix E. Model selection



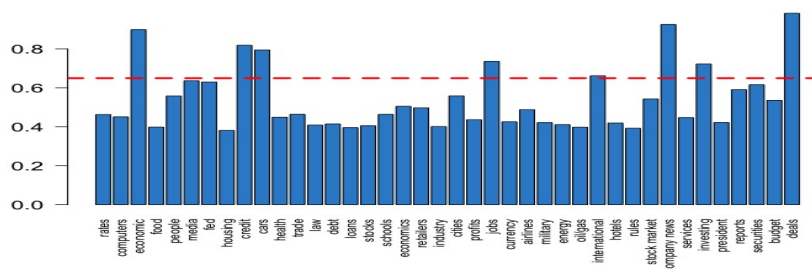
(a) Sign adjustment



(b) Non-adjusted frequency



(c) Uncertainty



(d) Positive tone for shocks from Gertler &amp; Karadi (2015)

Figure E.1: Posterior inclusion probabilities

Table E.1: LASSO and Elastic Net

	Lasso sign adj	Elastic net sign adj	Lasso tone	Elastic net tone	Lasso uncertainty	Elastic net uncertainty
Rates	-	-	-	-	-	-
Computers	-	-	-	-	-	-
Economic	0.001	0.001	0.002	0.002	-	-
Food	-	-	-	-	-	-
People	-	-	-	-	-	-
Media	-	-	-	-	-	-
Fed	-	-	-	-	-	-
Housing	-	-	-	-	-	-
Credit	-	-	0.002	0.002	-	-
Cars	-	-	-	-	-	-
Health	-	-	-	-	-	-
Trade	-	-	-	-	-	-
Law	-	-	-	-	-	-
Debt	-	-	-	-	-	-
Loans	-	-	-	-	-	-
Stocks	-	-	-	-	-	-
Schools	-	-	-	-	-	-
Economics	0.001	0.001	0.002	0.002	-	-
Retailers	-	-	-	-	-	-
Industry	-	-	-	-	0.004	0.004
Cities	-	-	-	-	-	-
Profits	-	-	-	-	-	-
Jobs	-	-	-	-	-	-
Currency	-	-	-	-	-	-
Airlines	-	-	-	-	-	-
Military	-	-	-	-	-	-
Energy	-	-	-	-	-	-
Oil/gas	-	-	-	-	-	-
International	-	-	0.002	0.002	-	-
Hotels	-	-	-	-	-	-
Rules	-	-	-	-	-	-
Stock market	-	-	-	-	-	-
Company news	-	-	0.001	0.001	-	-
Services	-	-	-	-	-	-
Investing	-	-	0.003	0.003	-	-
President	-	-	-	-	-	-
Reports	-	-	-	-	-	-
Securities	-	-	-	-	-	-
Budget	-	-	-	-	-	-
Deals	-	-	0.003	0.003	-	-

## Appendix F. OLS results

Table F.1: Positivity of Fed signals

	<i>Dependent variable:</i>				
	FFR_4G&K	FFR_4J&K	FFR_factorJ&K	SP500 J&K	SP500_factorJ&K
	(1)	(2)	(3)	(4)	(5)
Rates topic	-0.861 (3.342)	-1.572 (2.345)	-1.902 (3.278)	18.206 (19.400)	13.890 (15.586)
Economic topic	4.286*** (1.571)	3.994** (1.670)	4.809** (2.157)	-25.579* (14.599)	-19.515 (17.801)
Fed topic	-0.746*** (0.234)	-0.730*** (0.264)	-0.746** (0.353)	8.916** (3.805)	9.086** (3.583)
Credit topic	3.526 (3.913)	12.157*** (3.994)	15.879*** (5.303)	-109.175 (72.678)	-79.571 (65.686)
Debt topic	-1.001 (2.315)	-3.004 (2.288)	-0.606 (3.647)	-11.607 (28.171)	-14.921 (28.743)
Loans topic	1.153 (3.325)	5.015 (5.167)	4.468 (7.519)	-4.474 (61.407)	-12.950 (64.549)
Stocks topic	0.446 (3.746)	5.486** (2.494)	7.294** (3.542)	-18.581 (47.352)	1.641 (44.211)
Economics topic	0.002 (2.169)	2.191 (1.961)	2.314 (2.806)	5.517 (27.104)	-9.971 (22.708)
Jobs topic	-3.575*** (1.240)	-2.826*** (0.997)	-3.321** (1.573)	16.214 (14.567)	15.232 (17.930)
Currency topic	4.052 (2.902)	1.160 (2.584)	2.168 (3.437)	57.133 (45.594)	41.429 (44.132)
Energy topic	0.371 (2.627)	-3.999 (3.201)	-1.400 (4.317)	-64.982 (75.471)	-97.476 (76.015)
Oil/gas topic	-1.755 (1.398)	0.341 (2.137)	0.312 (2.684)	12.903 (16.938)	16.269 (16.081)
International topic	-2.411 (4.165)	3.328 (5.373)	2.969 (7.159)	79.164 (66.266)	33.057 (48.351)
Rules topic	0.390 (0.319)	0.055 (0.394)	-0.154 (0.606)	4.228 (4.873)	3.868 (4.603)
Stock market topic	-4.292 (2.921)	2.543 (4.056)	-0.230 (5.424)	12.624 (39.694)	12.854 (39.259)
Investing topic	8.324*** (2.765)	4.560* (2.341)	5.777** (2.698)	-32.853 (28.372)	-18.399 (21.088)
Reports topic	-3.500** (1.465)	-3.237** (1.443)	-3.740* (2.136)	15.280 (19.087)	9.000 (17.923)
Securities topic	-3.058 (3.285)	-0.002 (2.332)	-1.837 (3.279)	99.721*** (38.439)	85.937** (35.345)
Budget topic	-0.805 (0.964)	-1.024 (0.825)	-1.037 (1.070)	10.039 (9.691)	15.415 (13.330)
Deals topic	43.559*** (15.062)	35.250*** (10.358)	53.609*** (12.874)	-128.366 (114.284)	-128.047 (108.979)
Constant	-0.007** (0.003)	-0.006*** (0.002)	0.003 (0.003)	0.013 (0.030)	-0.004 (0.030)
AIC	-816.9	-938.4	-754.4	417.7	403.8
BIC	-742.3	-858.9	-675	497.2	483.3
Observations	220	274	274	274	274
R <sup>2</sup>	0.279	0.260	0.227	0.183	0.180
Adjusted R <sup>2</sup>	0.207	0.202	0.166	0.119	0.116

*Note:*

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Newey-West HAC standard errors are in parentheses

Table F.2: Uncertainty of Fed signals

	<i>Dependent variable:</i>				
	FFR_4G&K	FFR_4J&K	FFR_factorJ&K	SP500 J&K	SP500_factorJ&K
	(1)	(2)	(3)	(4)	(5)
Rates topic	-1.572 (1.991)	1.652 (3.059)	0.351 (3.850)	41.405 (42.034)	6.656 (36.910)
Economic topic	2.042* (1.138)	2.619* (1.389)	3.675** (1.854)	-46.019*** (15.847)	-41.499** (16.464)
Fed topic	0.709 (1.066)	1.001 (1.241)	1.171 (1.798)	25.950** (12.632)	24.784* (13.263)
Credit topic	-0.132 (2.362)	-0.561 (1.262)	0.571 (2.019)	-47.383** (23.066)	-29.990* (17.451)
Debt topic	0.676 (1.627)	-1.110 (2.097)	-1.459 (2.911)	18.711 (21.483)	2.106 (23.664)
Loans topic	2.625 (3.066)	-11.201 (8.681)	-15.323 (10.692)	95.078 (62.357)	107.442 (70.940)
Stocks topic	-4.921* (2.869)	-6.396 (4.995)	-9.286 (6.080)	59.032 (45.636)	64.884 (40.602)
Economics topic	-2.649 (2.020)	-4.403*** (1.679)	-4.715 (2.927)	29.360 (29.118)	58.364** (27.715)
Jobs topic	-5.301** (2.149)	1.616 (2.027)	3.590 (3.136)	-16.428 (30.447)	-23.850 (27.885)
Currency topic	-6.211** (2.515)	0.018 (5.140)	1.041 (7.235)	-34.483 (52.980)	-25.802 (60.214)
Energy topic	0.178 (3.343)	2.897 (3.426)	6.440 (5.285)	-7.778 (25.644)	-8.782 (27.870)
Oil/gas topic	1.905 (2.266)	1.970 (2.505)	0.414 (3.427)	-4.539 (24.677)	-1.006 (34.714)
International topic	-1.653 (3.419)	1.772 (3.710)	1.931 (5.249)	-72.030* (36.871)	-76.458* (44.361)
Rules topic	-3.450* (1.885)	-1.324 (1.590)	0.726 (2.498)	-39.405 (27.565)	-54.819** (26.898)
Stock market topic	0.340 (2.002)	0.911 (3.251)	4.216 (4.772)	-14.544 (41.281)	-26.083 (40.960)
Investing topic	-2.086 (1.633)	-0.225 (2.476)	2.369 (2.819)	-46.259* (25.559)	-24.975 (22.002)
Reports topic	1.920 (1.968)	-3.263 (2.285)	-4.267 (3.043)	22.774 (16.464)	22.072 (14.678)
Securities topic	2.258 (2.242)	2.818 (2.601)	4.798 (4.314)	37.428 (41.372)	22.339 (29.652)
Budget topic	-0.832 (2.010)	2.223 (4.482)	1.222 (6.011)	6.548 (40.933)	-6.518 (34.723)
Deals topic	-30.786*** (6.569)	-15.786*** (6.055)	-20.043** (8.991)	33.463 (75.676)	45.482 (85.484)
Constant	-0.008** (0.003)	-0.007** (0.003)	0.003 (0.004)	0.011 (0.029)	-0.001 (0.027)
AIC	-774.8	-887	-713.5	435.1	420.6
BIC	-700.1	-807.5	-634	514.6	500.1
Observations	220	274	274	274	274
R <sup>2</sup>	0.127	0.108	0.102	0.130	0.129
Adjusted R <sup>2</sup>	0.039	0.037	0.031	0.061	0.060

*Note:*

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Newey-West HAC standard errors are in parentheses

Table F.3: Positivity of Fed signals. All variables

	<i>Dependent variable:</i>				
	FFR_4G&K	FFR_4J&K	FFR_factorJ&K	SP500 J&K	SP500_factorJ&K
	(1)	(2)	(3)	(4)	(5)
Rates topic	-2.173 (2.730)	-2.801 (2.275)	-3.844 (3.712)	24.472 (20.800)	24.066 (19.292)
Computers topic	-25.390*** (4.131)	-11.180** (5.404)	-15.619* (8.761)	139.830* (80.715)	222.257*** (76.790)
Economic topic	4.091*** (1.401)	3.471** (1.581)	4.303** (2.171)	-17.996 (18.040)	-12.152 (18.472)
Food topic	7.787 (10.477)	7.972** (3.813)	8.191 (6.226)	72.343 (50.150)	57.029 (39.834)
People topic	-2.728 (6.470)	-3.296 (8.881)	-8.110 (12.645)	112.226 (89.387)	108.205 (89.379)
Media topic	6.748 (9.029)	9.505 (7.090)	16.143** (8.152)	21.095 (127.036)	63.638 (88.792)
Fed topic	-0.623** (0.246)	-0.613** (0.256)	-0.566* (0.340)	7.196** (3.276)	6.974** (2.800)
Housing topic	0.179 (1.078)	1.400 (1.022)	0.895 (1.164)	-13.710 (11.148)	-15.805 (10.330)
Credit topic	-4.631 (3.544)	13.468*** (4.734)	13.992* (7.728)	-48.848 (53.116)	-6.348 (47.047)
Cars topic	-3.682 (5.839)	-26.085*** (8.295)	-27.396** (11.556)	74.033 (93.481)	74.970 (94.416)
Health topic	18.368** (7.328)	9.962* (5.268)	10.545 (7.441)	-189.398** (84.425)	-156.862** (67.408)
Trade topic	4.156 (4.760)	0.675 (4.191)	0.886 (5.590)	-74.544* (45.249)	-70.026* (42.228)
Law topic	-4.124 (3.525)	-2.238 (4.575)	-2.450 (5.998)	-16.306 (45.636)	-4.978 (49.408)
Debt topic	-0.946 (2.500)	-4.168 (3.464)	-1.509 (5.011)	-6.523 (36.752)	-20.337 (34.272)
Loans topic	0.619 (3.009)	4.001 (5.900)	4.333 (9.113)	-12.977 (58.126)	-14.571 (58.822)
Stocks topic	0.781 (2.796)	5.940** (2.635)	7.353** (3.739)	-21.030 (40.772)	6.217 (31.551)
Schools topic	-16.287 (12.721)	-9.899 (19.204)	-0.518 (29.912)	32.554 (292.621)	87.252 (312.142)
Economics topic	-2.414 (1.941)	1.580 (2.147)	1.106 (2.985)	5.416 (28.453)	-6.885 (25.594)
Retailers topic	-10.120* (5.673)	-5.882 (5.135)	-10.803 (7.482)	-11.925 (75.950)	-0.843 (68.778)
Industry topic	4.060* (2.117)	-1.107 (4.484)	-1.572 (6.590)	-24.745 (57.620)	-32.693 (69.568)
AIC	-825.1	-931	-743.4	398.6	375.8
BIC	-682.6	-779.2	-591.6	550.3	527.5
Observations	220	274	274	274	274
R <sup>2</sup>	0.421	0.343	0.304	0.342	0.361
Adjusted R <sup>2</sup>	0.292	0.231	0.185	0.229	0.251

*Note:*

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Newey-West HAC standard errors are in parentheses

Errors from the first and forth model are autocorrelated at 10% level



Table F.4: Positivity of Fed signals. All variables

	<i>Dependent variable:</i>				
	FFR_4G&K	FFR_4J&K	FFR_factorJ&K	SP500 J&K	SP500_factorJ&K
	(1)	(2)	(3)	(4)	(5)
Cities topic	-8.349 (5.217)	-2.859 (5.001)	-1.644 (8.217)	3.173 (78.301)	-32.484 (53.957)
Profits topic	-0.681 (1.982)	0.427 (3.221)	-0.296 (3.714)	29.476 (22.997)	30.422 (27.137)
Jobs topic	-4.457*** (1.527)	-2.362* (1.431)	-2.797 (2.056)	2.912 (18.522)	0.466 (16.055)
Currency topic	6.449** (3.146)	5.935 (3.829)	8.561 (5.402)	-10.254 (50.261)	-34.083 (49.558)
Airlines topic	-22.822 (14.549)	-53.283*** (19.599)	-60.975*** (22.286)	194.673 (368.255)	161.224 (451.712)
Military topic	2.112 (5.130)	-7.075 (8.323)	-6.121 (8.222)	34.323 (69.534)	53.086 (72.473)
Energy topic	-3.604 (2.537)	-5.224 (3.376)	-4.521 (4.278)	2.747 (54.385)	-23.479 (46.715)
Oil/gas topic	-1.459 (1.094)	-0.506 (1.822)	-0.928 (2.438)	19.871 (19.841)	24.849 (17.518)
International topic	-3.159 (4.052)	7.252 (7.036)	5.658 (8.783)	46.580 (63.494)	10.586 (56.792)
Hotels topic	5.845 (8.936)	8.457 (11.137)	-18.880 (27.486)	33.143 (192.150)	50.952 (216.967)
Rules topic	0.459 (0.405)	-0.043 (0.490)	-0.304 (0.668)	2.748 (5.457)	1.680 (5.220)
Stock market topic	-5.622** (2.710)	5.685 (3.795)	3.658 (4.803)	17.182 (39.350)	11.360 (42.972)
Company news topic	9.654** (4.336)	10.734 (7.588)	13.320 (9.026)	-191.549*** (68.725)	-187.719** (73.896)
Services topic	24.698*** (7.556)	12.950* (7.668)	9.365 (14.106)	-90.610 (137.219)	-154.051 (144.631)
Investing topic	6.627*** (1.956)	3.755 (2.693)	5.324* (3.148)	-5.570 (34.199)	7.979 (26.456)
President topic	8.266*** (2.655)	3.110 (2.449)	2.923 (3.643)	-23.821 (32.145)	-23.008 (26.880)
Reports topic	-1.863 (1.293)	-3.128** (1.385)	-2.921 (1.873)	12.165 (16.198)	4.390 (15.359)
Securities topic	5.154 (3.343)	1.632 (2.560)	1.088 (3.766)	75.634*** (28.592)	69.571** (27.949)
Budget topic	-0.933 (0.890)	-1.639** (0.668)	-1.742* (0.909)	11.254* (6.645)	17.845** (7.730)
Deals topic	35.983*** (10.505)	41.247*** (12.851)	64.656*** (15.373)	-149.433 (114.130)	-191.419* (99.468)
Constant	-0.004* (0.003)	-0.007*** (0.002)	0.002 (0.003)	-0.005 (0.028)	-0.024 (0.027)
AIC	-825.1	-931	-743.4	398.6	375.8
BIC	-682.6	-779.2	-591.6	550.3	527.5
Observations	220	274	274	274	274
R <sup>2</sup>	0.421	0.343	0.304	0.342	0.361
Adjusted R <sup>2</sup>	0.292	0.231	0.185	0.229	0.251

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

Newey-West HAC standard errors are in parentheses  
Errors from the first and forth model are autocorrelated at 10% level

Table F.5: Sign adjustment of Fed signals

	<i>Dependent variable:</i>				
	FFR_4G&K	FFR_4J&K	FFR_factorJ&K	SP500 J&K	SP500_factorJ&K
	(1)	(2)	(3)	(4)	(5)
Rates topic	0.143 (0.152)	-0.163 (0.141)	-0.219 (0.213)	1.748 (1.249)	1.191 (1.020)
Economic topic	0.242** (0.105)	0.218** (0.098)	0.265** (0.114)	-1.377 (0.862)	-1.009 (0.961)
Fed topic	-0.082*** (0.026)	-0.045** (0.022)	-0.042 (0.028)	0.686** (0.348)	0.672** (0.312)
Credit topic	0.119 (0.247)	0.368** (0.166)	0.371 (0.240)	-0.061 (3.729)	-0.580 (3.880)
Debt topic	-0.317* (0.189)	-0.104 (0.160)	0.071 (0.216)	-2.085 (1.886)	-2.062 (1.927)
Loans topic	-0.063 (0.147)	-0.133 (0.179)	-0.280 (0.240)	3.441 (2.941)	2.373 (2.386)
Stocks topic	0.135 (0.243)	0.372* (0.210)	0.518** (0.264)	-1.865 (2.302)	-0.404 (2.252)
Economics topic	0.192* (0.101)	0.062 (0.110)	0.094 (0.140)	0.385 (1.474)	-0.602 (1.513)
Jobs topic	-0.206* (0.113)	-0.174** (0.075)	-0.272** (0.118)	2.152** (1.037)	2.036* (1.056)
Currency topic	0.268 (0.192)	0.348* (0.180)	0.556** (0.267)	0.248 (2.424)	-0.548 (2.274)
Energy topic	-0.051 (0.142)	-0.050 (0.123)	0.057 (0.167)	-5.120 (4.119)	-5.967 (3.652)
Oil/gas topic	-0.167 (0.106)	0.020 (0.110)	-0.020 (0.134)	0.944 (1.170)	1.083 (1.078)
International topic	-0.010 (0.200)	0.095 (0.246)	-0.032 (0.307)	4.508 (3.118)	2.288 (2.392)
Rules topic	0.070 (0.049)	0.022 (0.040)	0.052 (0.049)	-0.338 (0.351)	-0.202 (0.394)
Stock market topic	-0.199 (0.178)	0.143 (0.292)	-0.004 (0.381)	-1.149 (3.048)	-1.336 (2.869)
Investing topic	0.484*** (0.164)	0.303** (0.143)	0.390** (0.175)	-2.637* (1.494)	-1.946 (1.278)
Reports topic	-0.137 (0.086)	-0.197** (0.085)	-0.220* (0.125)	0.786 (0.980)	0.466 (1.027)
Securities topic	-0.075 (0.173)	0.009 (0.153)	-0.045 (0.228)	5.268** (2.439)	4.934** (2.384)
Budget topic	-0.107 (0.155)	-0.068 (0.058)	-0.115 (0.100)	1.552 (1.348)	1.670 (1.460)
Deals topic	1.180*** (0.419)	1.355*** (0.461)	1.933*** (0.543)	-10.463** (5.011)	-9.381** (3.918)
Constant	-0.007*** (0.003)	-0.006** (0.002)	0.003 (0.003)	0.006 (0.033)	-0.009 (0.032)
AIC	-820	-918.5	-739.2	421.8	407.4
BIC	-745.3	-839	-659.7	501.3	486.9
Observations	220	274	274	274	274
R <sup>2</sup>	0.289	0.205	0.182	0.171	0.170
Adjusted R <sup>2</sup>	0.218	0.142	0.118	0.106	0.104

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Newey-West HAC standard errors are in parentheses

Table F.6: Frequency of Fed signals

	<i>Dependent variable:</i>				
	FFR_4G&K	FFR_4J&K	FFR_factorJ&K	SP500 J&K	SP500_factorJ&K
	(1)	(2)	(3)	(4)	(5)
Rates topic	-0.071 (0.086)	0.161 (0.117)	0.276* (0.154)	0.497 (1.265)	0.612 (1.262)
Economic topic	-0.035 (0.049)	0.017 (0.055)	0.038 (0.077)	-1.165* (0.606)	-0.955 (0.638)
Fed topic	0.071* (0.038)	0.009 (0.028)	-0.001 (0.037)	-0.128 (0.488)	0.282 (0.427)
Credit topic	-0.178* (0.093)	-0.207* (0.115)	-0.156 (0.153)	-1.545 (1.453)	-0.598 (1.430)
Debt topic	-0.015 (0.090)	0.069 (0.119)	0.158 (0.158)	-2.724** (1.178)	-2.809** (1.166)
Loans topic	0.045 (0.071)	0.041 (0.117)	-0.067 (0.135)	0.782 (1.167)	0.271 (1.319)
Stocks topic	0.034 (0.108)	-0.035 (0.147)	-0.127 (0.205)	3.571 (2.213)	3.069 (2.276)
Economics topic	-0.002 (0.056)	0.025 (0.043)	0.006 (0.076)	0.373 (0.740)	0.789 (0.774)
Jobs topic	-0.060 (0.106)	0.112 (0.078)	0.134 (0.093)	-1.150 (1.103)	-0.982 (1.109)
Currency topic	-0.102 (0.103)	0.039 (0.109)	0.076 (0.129)	-3.432* (1.857)	-2.298 (1.795)
Energy topic	-0.104 (0.099)	0.087 (0.140)	0.035 (0.197)	1.540 (1.488)	1.093 (1.323)
Oil/gas topic	0.080 (0.055)	0.029 (0.074)	0.036 (0.114)	-0.745 (1.267)	-0.081 (1.187)
International topic	0.107 (0.086)	0.234 (0.203)	0.382 (0.269)	-2.636 (1.929)	-1.397 (2.132)
Rules topic	-0.122 (0.082)	-0.133** (0.065)	-0.118 (0.093)	0.854 (0.759)	0.620 (0.738)
Stock market topic	0.184* (0.095)	0.120 (0.101)	0.243 (0.156)	-0.199 (1.085)	-0.134 (1.075)
Investing topic	-0.106 (0.069)	-0.123 (0.136)	-0.164 (0.188)	0.518 (1.167)	0.850 (1.147)
Reports topic	0.149** (0.067)	0.057 (0.069)	0.075 (0.096)	0.449 (0.783)	0.519 (0.691)
Securities topic	-0.073 (0.053)	-0.009 (0.062)	0.008 (0.097)	0.281 (1.057)	0.714 (1.054)
Budget topic	-0.097 (0.103)	-0.059 (0.082)	-0.111 (0.132)	2.367* (1.217)	2.557** (1.246)
Deals topic	-0.211 (0.220)	-0.030 (0.209)	-0.254 (0.298)	2.783 (1.741)	3.174** (1.619)
Constant	-0.008** (0.004)	-0.007** (0.003)	0.003 (0.004)	0.009 (0.026)	-0.005 (0.025)
AIC	-769.1	-889.3	-714.8	409	405
BIC	-694.4	-809.8	-635.3	488.5	484.5
Observations	220	274	274	274	274
R <sup>2</sup>	0.104	0.115	0.106	0.209	0.177
Adjusted R <sup>2</sup>	0.014	0.045	0.036	0.147	0.112

*Note:*

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Newey-West HAC standard errors are in parentheses

## Appendix G. The Bayesian Vector Autoregression

I use Bayesian Vector Autoregression (BVAR) with an independent normal-inverted Wishart prior for the reduced form coefficients (see Koop & Korobilis (2010) for more details):

$$\begin{aligned} p(\beta, Q) &= p(\beta)p(Q) \\ p(\beta) &\sim f_N(\beta|\underline{\beta}, \underline{V}_\beta) \\ p(Q) &\sim f_{IW}(Q|\underline{Q}, \underline{v}_Q) \end{aligned}$$

For dealing with overfitting I entertain a prior in Minnesota fashion. Prior for  $\beta_m$  (3-month federal funds futures and S&P 500 surprises) is set to 0, other  $\beta$  at 1 for its own lags, and zero everywhere else.  $\underline{V}_\beta$  is a diagonal matrix implying that the standard deviation of lag  $l$  of variable  $j$  in equation  $i$  is  $\frac{\lambda_1 \lambda_2 \sigma_i}{\sigma_j l \lambda_3}$  for  $j \neq i$ ,  $\frac{\lambda_1}{l \lambda_3}$  for  $j = i$  and  $\lambda_4 \sigma_i$  for a constant. I use standard hyperparameters from the literature:  $\lambda_1 = 0.2$ ,  $\lambda_2 = 0.5$ ,  $\lambda_3 = 1$ ,  $\lambda_4 = 100$ .  $\sigma_i, \sigma_j$  are scaled measures of the variance associated with the AR(p) equation estimate.  $\underline{Q}$  is a diagonal matrix. Lastly, I set  $\underline{v}_Q = 10$ . Based on the priors the conditional posterior for  $\beta$  is:

$$\begin{aligned} \beta|y, Q^{-1} &\sim N(\bar{\beta}, \bar{V}_b) \\ \bar{V}_b &= (\underline{V}_\beta^{-1} + \sum_{t=1}^T X_t' Q^{-1} X_t)^{-1} \\ \bar{V}_b &= \bar{V}_b (\underline{V}_\beta^{-1} \underline{\beta} + \sum_{t=1}^T X_t' Q^{-1} y_t) \end{aligned}$$

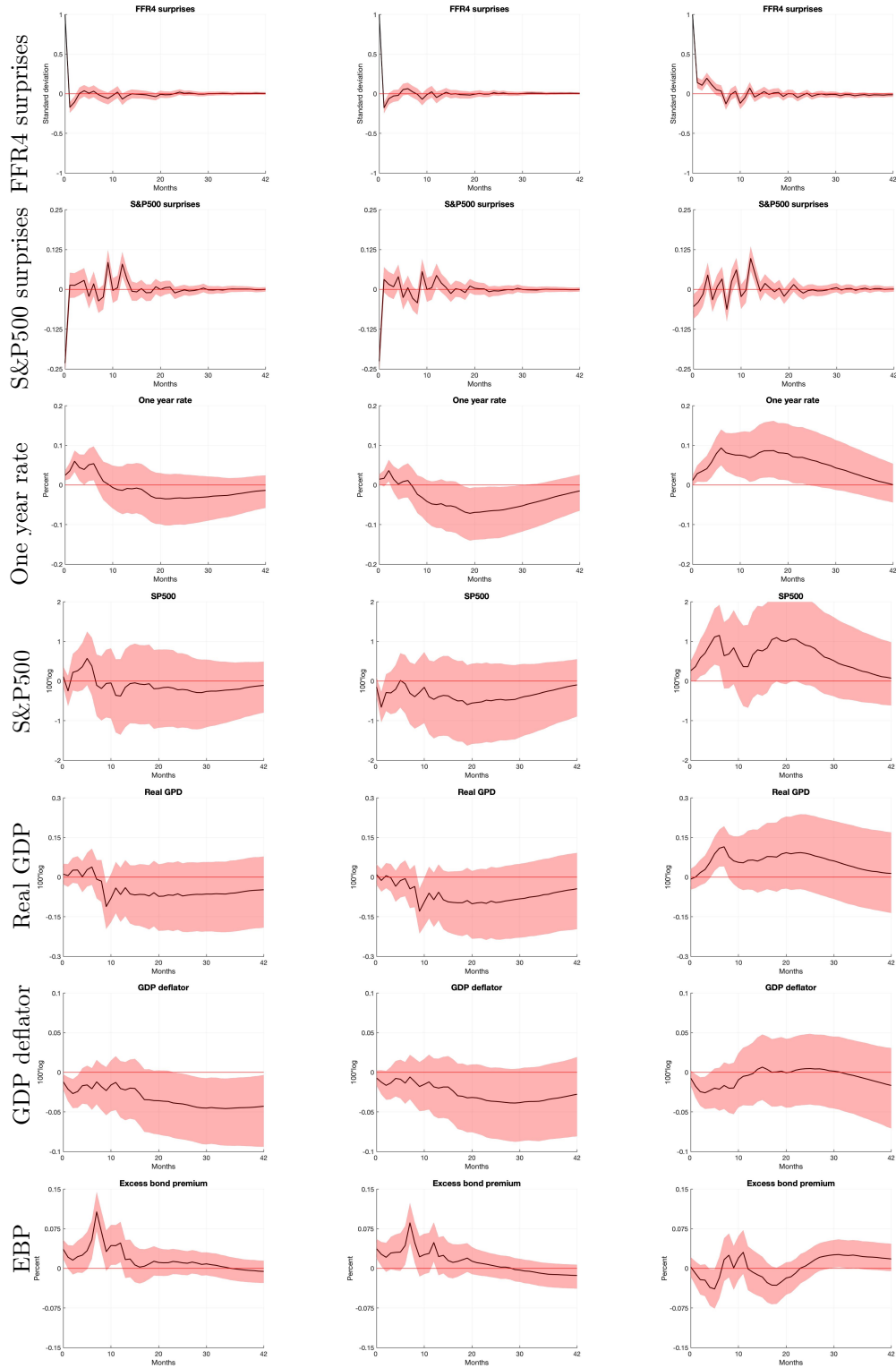
The conditional posterior of  $Q$  is:

$$\begin{aligned} Q|y, \beta &\sim IW(\bar{Q}, \bar{v}_Q) \\ \bar{v}_Q &= \underline{v}_Q + T \\ \bar{Q} &= \underline{Q} + \sum_{t=1}^T (y_t - X_t' \beta)(y_t - X_t' \beta)' \end{aligned}$$

12,000 Gibbs sampler draws were taken in total and 2,000 were discarded after burn-in.

The SVAR has 12 lags. The sample is monthly, from March 1994 to December 2016.

# Appendix H. Results for sign adjusted topics in FOMC statements



(a) Baseline

(b) Policy shock

(c) Information shock

Figure H.1: Comparison between monetary policy and information shocks shaded 16% and 84% percentiles