MPRA
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/104604/
MPRA Paper No. 104604, posted 13 Dec 2020 20:55 UTC

Central Bank Communication: Information and Policy shocks *

Nataliia Ostapenko †

First version: May 2020
This version: December 6, 2020

Abstract

The study proposes a novel way to identify the effects of monetary policy shocks taking into account time-varying signals of the central bank. I augment the standard monetary policy Bayesian Vector Autoregression (BVAR) with additional information variables from Fed statements, which allows us to study the information-free effects of monetary policy shocks and to take into account forward-looking information released by the central bank. The results show that, compared to surprises in 3-month federal funds futures, the policy shock identified in this study has a more negative effect on GDP, a more prolonged negative effect on inflation and a greater impact effect on the excess bond premium. In the short-run it causes S&P500 to decline and the Fed to raise its interest rate. Furthermore, the results of large-scale Bayesian VAR confirm the standard transmission channels of monetary policy.

Keywords: monetary policy, shock, transmission, statements, Latent Dirichlet Allocation, information

JEL Classification: E52, E31, E00

*I am grateful to Eleonora Granziera, Alessia Paccagnini, Lenno Uusküla, and to the participants of 2020 Central Bank Macro Modelling Workshop for helpful comments and suggestions. Any errors that remain are my own.

†University of Tartu. Contact: nataliia.ostapenko@gmail.com

1 Introduction

How does monetary policy affect the economy? To answer this question it is necessary to find an effective means of analysing the effects of monetary policy shocks due to the fact that the Fed reacts to macroeconomic indicators and shocks should be orthogonal to this reaction. The principal empirical strategies lie in purging a monetary policy instrument from the reaction function (Romer & Romer (2004)) or employing high-frequency identification (Gertler & Karadi (2015)). Nevertheless, recent studies have pointed out that the effect of the information in central bank communications 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 the Federal Open Market Committee (FOMC) might possess insider information (Romer & Romer (2000)), as a consequence of which FOMC statements might release this private information or time-varying preferences of the central bank to the public. The reaction in a narrow window around Fed announcement could well contain a response to this additional information rather than a response to unexpected monetary policy action by the Fed. Therefore, the 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 (forthcoming) found that shocks identified by purging can be predicted from macroeconomic indicators (from *Federal Reserve Economic Data* (2019)), while shocks identified by high-frequency strategy can be predicted from *Greenbook Historical and Forecast Data* (2019) projections.

Moreover, literature still lacks a good measure of monetary policy shock. Popular approaches rely on purging suitable proxies with respect to Fed private forecasts (Romer & Romer (2004), Miranda-Agrippino & Ricco (forthcoming)) or on augmenting VAR with forward-looking information emanating from Federal Reserve forecasts (Bu et al. (2020)). Nevertheless, these approaches are also problematic because Fed forecasts are not available to the public in real-time (as was pointed out by Ramey (2016)) and, therefore, the "correct" reactions of macroeconomic variables to a "monetary policy shock" in this case are puzzling because there is still unresolved signal-extraction problem by the public. Using information released in policy statements instead of Fed private forecasts helps to overcome above-mentioned problems and at the same time to purge monetary policy surprises.

This paper provides original empirical evidence regarding the information contained in FOMC statements. Identifying the type of information inherent in “policy surprises” enables these surprises to be decomposed into information and information-free policy effects. I use FOMC statements as the main data source for 1994–2016 because the Fed started to release statements from 1994. I employ Latent Dirichlet Allocation (LDA) pre-trained on the business sections of main U.S. newspapers for content extraction from the FOMC statements, which transforms Fed statements into topics distributions over time. Following that, I adopt a lexical-based approach to assign the tone to each sentence from the FOMC statements, which counts the frequency 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 the 3-month federal funds futures on FOMC statement release dates. I use Bayesian Lasso regression for this purpose.

Furthermore, the study decomposes federal funds future surprises on FOMC dates into information and information-free shocks by augmenting the standard VAR with information variables, which enable us to separate 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 futures and S&P500 surprises in a narrow window around the FOMC announcement, as well as the main macroeconomic indicators employed in Vector Autoregressions (VAR) by Jarocinski & Karadi (2020) to make the findings comparable.

The main results are as follows. The most important news released by the central bank concern the macroeconomy. The positivity of these signals lead to an increase in short-term nominal daily yields, while signals concerning macroeconomic uncertainty increase long-term daily nominal and real yields, as well as expected inflation. These fiddlings are also in line with those of Hansen et al. (2019). The result confirms that central bank communication is a multi-dimensional object and affects the economy in different directions.

The conventional way of identifying monetary policy shocks in the literature is to rotate principal components around monetary policy announcements to capture target and path factors of monetary policy, as was first introduced by the excellent work of Gürkaynak et al. (2005). Derived path factor should capture communication channel of monetary policy that influence long-term rates mainly through the term premium. My findings show that this path factor might capture central bank uncertain signals concerning future economic development. These signals are important for the long tails of nominal and real yield curves.

The paper introduces a novel way to identify the effects of monetary policy shocks conditional on information released by the central bank in its statements. The popular approach is to purge suitable surprise components with respect to Fed private forecasts (Romer & Romer (2004), Miranda-Agrippino & Ricco (forthcoming)). But this approach has two major issues. Firstly, why to do conditioning on Fed private forecasts that are unavailable to the public in real-time? In this case, there is still a signal extraction problem by the public and the "standard" reactions of output and inflation are questionable[1]. Secondly and more importantly, recent literature questions the information advantages of the Fed[2]. An alternative explanation is that the Fed releases its time-varying reaction function to the public and from its statements the public learns it and updates expectations accordingly. In this case, it is reasonable to conditioning on central bank signals[3] instead of internal forecasts. Moreover, Gürkaynak et al. (2020) shows that based on the comments in the financial press, latent factors that explain most of the variation of the yield curve are indeed days of well-known "statement surprises".

After augmenting VAR with these news series 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. What is more, it rises the cost of credit on impact.

The results contribute to the literature on the transmission channel of monetary policy. The results of large-scale Bayesian VAR show that a monetary policy shock is transmitted according to the theory: it reduces real economic activity, inflation, credit spreads, while increases interest rates, the cost of credit and macroeconomic uncertainty. The results also confirm the importance of interest rate, credit, exchange rate, asset prices and expectations channels of monetary policy propagation. However, contrary to previous findings, I could not confirm the importance of the term premium channel for monetary policy shocks propagation.

The findings add to the results of Jarocinski & Karadi (2020), who employed sign restrictions to identify the monetary policy and information effects of the central bank. The effect of policy surprise shocks are in line with the main findings of Jarocinski & Karadi (2020): the effect is less persistent on interest rates but more persistent on inflation

[1] As was pointed out by Ramey (2016)

[2] See, for instance, Michael D. Bauer and Eric T. Swanson (2020)

[3] Delphic forward guidance by Campbell et al. (2012)

and the cost of credit. Therefore, I relaxed sign restrictions and obtained quantitatively similar results.

Moreover, the study extends the findings of Romer & Romer (2000) on asymmetric information between the Federal Reserve and the public. My findings show that FOMC statements provide additional information which goes beyond the monetary policy actions themselves, but this information also should be well anticipated by markets. Therefore, the central bank might release time-varying policy preferences to the public in its communications instead of new information about economic development.

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. Hubert (2019) found that contractionary monetary policy has negative effects on inflation expectations and stock prices only if and when associated with positive economic news. This study could not confirm this finding. Moreover, Iglesias et al. (2017) found that neither positive nor negative communication had particularly significant effects on inflation nor real economic activity, whereas Hubert & Labondance (2017) found that sentiment affects private interest rate expectations, inflation and industrial production beyond monetary shocks. On the contrary, this study finds that communication mainly reduces the cost of credit.

The study expands the literature on the importance of FOMC statements. For instance, Gürkaynak et al. (2020) show that when a sample includes statements the heteroskedasticity-based estimator yields a reaction coefficient that is two to 400 times larger than the OLS estimator without statements.

Last but not least, this study complements the recent literature in its way of decomposing FOMC statements into topic time series with sentiments. To the author's best knowledge this is the first study to employ a 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. Subsequently, the previously-cited authors assigned the tone to each topic. My approach differs from the above-mentioned in that the LDA model was trained on the business section of a selection of U.S. newspapers, which enables us to obtain more distinguishable topics. What is more, my methodology captures changes in the topic composition of FOMC statements without the need to rely on the dynamic topic model.

The remainder of the paper proceeds as follows. Section 2 describes the data and methodology. Section 3 discusses the information content of Fed communication. Section 4 discusses the mechanism of central bank communication effects. Section 5 presents the main results on the transmission mechanism of the information-free effect of monetary policy. Section 6 concludes.

2 Methodology

The Federal Open Market Committee (FOMC) holds eight regularly scheduled meetings during the year and additional meetings as needed. In these meetings the Federal Open Market Committee decides on the interest rate changes necessary 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 used in this study were downloaded from the Fed webpage.

The standard high-frequency identification strategy employs a narrow window (30 minutes) in order to detect surprise changes in 3-month federal funds futures around FOMC announcements. The main concern with this identification strategy lies in the fact that the FOMC might possess insider information (Romer & Romer (2000)), and FOMC announcements might contain additional information for the public or it might reveal its time-varying preferences. The reaction in a narrow window might contain a response to this additional information rather than 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 previous logic, Miranda-Agrippino & Ricco (forthcoming) found that the surprises highlighted in Gertler & Karadi (2015) can be predicted from from *Greenbook Historical and Forecast Data* (2019) projections and *Federal Reserve Economic Data* (2019) factors. The authors purged the shock series with respect to their own lags and Greenbook information[4]. Nevertheless, these surprises may, in fact, be attributed to revelation of time-varying preference of the central bank. In this case, purging shock series with respect to

[4]As in Romer & Romer (2004)

Greenbook information might not capture this additional information. Therefore, I purge shock series with respect to topics from FOMC statements. These topics and the tone of the Fed should serve to capture the Fed information effect and allows to separate 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. However, in order to obtain more accurate results when purging surprises in federal funds futures from the information effect, I add additional informational proxies to the standard SVAR, which allows us to separate these surprises from the effects of information.

To train a model for the topic extraction, details of which 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) and The Chicago Tribune (1985–2019). The New York Times is the second-largest in circulation and the largest circulating metropolitan newspaper with a weekly circulation of 2.1 million. It is also ranked 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 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 it 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, such as 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 $k = 1, ..., K$ (K is the total number of latent topics):
   - A discrete probability distribution over a fixed vocabulary that represents the k^{th} topic distribution, $\varphi_k \sim Dirichlet(\beta)$[5]
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), $\theta_d \sim Dirichlet(\alpha)$[6]
   - 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 $w_{d,n}$), $z_{d,n} \sim Mult(\theta_d)$[7]
     (b) An observed word, $w_{d,n} \sim Mult(\varphi_k)$

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

$$p(w_{d,n}, z_{d,n}, \theta_d, \varphi_k|\alpha, \beta) = (\prod_{n=1}^{N} p(z_{d,n}|\theta_d)p(w_{d,n}|z_{d,n}, \varphi_{n,k}))(\prod_{k=1}^{K} p(\varphi_k|\beta))(\prod_{d=1}^{D} p(\theta_d|\alpha))$$

$$= (\prod_{n=1}^{N} Mult(z_{d,n}|\theta_d)Mult(w_{d,n}|z_{d,n}, \varphi_{d,k}))(\prod_{k=1}^{K} Dirichlet(\varphi_k|\beta))(\prod_{d=1}^{D} Dirichlet(\theta_d|\alpha)) \quad (1)$$

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

Latent variables $z_{d,n}$, θ_d, φ_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 θ_d and φ_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 θ_d and word distributions per topic φ_k.

---

[5] $Dirichlet(.)$ is the Dirichlet distribution (a conjugate prior for the Multinomial distribution), β is a hyper-parameter

[6] α is a hyper-parameter.

[7] $Mult(.)$ is the Multinomial distribution.

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. However, 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 the statement's topic and its 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 with regard to labelling sentiment.

The positive sentiment of a sentence is calculated as following (2):

$$Pos_i = \frac{\#positivewords_i - \#negativewords_i}{\#totalwords_i} \tag{2}$$

The total monthly positive sentiment for a specific economic topic is calculated as the sum of sentence positive sentiments minus negative sentiments multiplied by topic proportions within the sentence and sum over the sentences (3):

$$Pos_{topic} = \sum_{i \in topic} Pos_i \times topic\_proportions_i \tag{3}$$

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).

Similarly, I calculated uncertainty sentiments by employing (2) and (3) for uncertain words from Loughran & Mcdonald (2016)[8].

In order to purge monetary policy shock series with respect to central bank signals I augment the standard VAR with information variables.

The standard Structural VAR representation in companion form is:

$$Ay_t = Bx_{t-1} + \varepsilon_t \tag{4}$$

[8] The full list of words for each sentiment category is available at https://sraf.nd.edu/textual-analysis/resources/

, where $x_t' = [1, y_{t-1}', y_{t-2}', ..., y_{t-k}']$ and $\varepsilon_t \sim i.i.d.N(0, D)$, D is diagonal, y_t is $k \times 1$ vector of endogenous variables, ε_t is $k \times 1$ vector of exogenous random shocks. A and B are $k \times k$ coefficient matrices.

The reduced form representation is:

$$y_t = A^{-1}Bx_{t-1} + A^{-1}\varepsilon_t$$

$$y_t = Cx_{t-1} + \epsilon_t \tag{5}$$

where $C = A^{-1}B$ and ϵ are $k \times 1$ vector of reduced form shocks, which don't have any economic interpretation.

Infinite MA representation of (5) is:

$$y_t = \sum_{j=0}^{\infty} \Theta_j \varepsilon_{t-j} \tag{6}$$

The identification problem to (6) comes from two sources. First one is the regular identification problem: recovering structural shocks from (6) by imposing restrictions on A matrix and variance-covariance matrix of residuals. These helps to overcome the observation equivalent problem of:

$$y_t = \sum_{j=0}^{\infty} (\Theta_j U^{-1})(U\varepsilon_{t-j})$$

$$y_t = \sum_{j=0}^{\infty} \Theta_j^* \varepsilon_{t-j}^* \tag{7}$$

When decision's maker information set is different from econometrician information set the second problem of non-uniqueness arises:

$$H^*(z)E\varepsilon^*\varepsilon^{*'}H^*(z^{-1})' = H(z)E\varepsilon\varepsilon'H(z^{-1})' \tag{8}$$

where $H(z)$ is the z-transform. The (8) shows two observatory equivalent results, one of which is invertible representations and other is non-invertible. If A^{-1} is equal to $H^*(z)$ then standard identification by imposing restrictions in A matrix recovers ε^* structural shock:

$$y_t = A^{-1}Bx_{t-1} + A^{-1}\varepsilon_t^*$$

Identifying A^{-1} recovers the shocks ε_t^*, but not the structural shocks ε_t, that agents observe since the econometrician conditions on a smaller information set than do agents (Leeper et al. (2013)). Moreover, there should not be foresight effects in VAR. Therefore, augmenting VAR with additional information variables that are forward-looking should help to overcome the invertibility problem of VAR. What is more, these information variables are available to the public in real-time and it is more reasonable to take this information into account instead of conditioning on the Fed information set that is not available to the public.

Noh (2018) suggested to use proxy variables as additional regressors in the VAR instead of using proxy variables as IV for a shock identification assuming the invertibility condition, because the Proxy-SVAR, which is the most efficient approach under the invertibility and linearity, is valid if and only if the pre-whitened proxy variable has no direct forecasting ability if it is added in the VAR. It is well-known that surprises in 3-m federal funds futures on FOMC announcements dates contain forward guidance effects and, therefore, have forecasting power for future interest rate changes. That is why instead of using surprises in 3-m federal funds futures as a proxy variable and assume invertibility of VAR, I use it as an additional regressor in the conventional monetary VAR. This is "internal instrument" approach, also pointed out by Plagborg-Mollerand & Wolf (2019), who highlighted that structural estimation with an instrument (proxy) can be carried out by ordering the instrument first in a recursive VAR, even under non-invertibility[9].

There are some implicit assumptions while using surprises around FOMC announcements to measure the effect of monetary policy shocks, namely (1) there is only one event in a selected window; (2) markets know exactly data-generating process and information of the central bank, (3) markets know exactly the central bank reaction function; (4) efficient market hypothesis; (5) a risk premium does not change in a window. Moreover, foresight should be already taken into account by markets. In this case, asset price changes in a window around an announcement can be represented as:

$$p_t^h - p_{t-30min}^h = [\mathbb{E}_t(i_{t+h}) - \mathbb{E}_{t-30min}(i_{t+h})] + [\zeta_t - \zeta_{t-30min}]$$

$$p_t^h - p_{t-30min}^h = [\mathbb{E}_t(i_{t+h}) - \mathbb{E}_{t-30min}(i_{t+h})] = e_p + error$$

where the first part in brackets is a revision of expectations and the second is a revision of a term premium.

A shock is an innovation orthogonal to the state of the economy and a surprise is an innovation orthogonal to the public information set. In case the Fed has an information

[9]For instance, Durante et al. (2020) used poor man's proxy of surprises as a policy shock measure in the framework of Jorda local projections.

advantage over markets, agents' update of forecasts during an announcement can confound e_p with the reaction function of the central bank, which gives a reason to purge surprises with respect to the central bank information set. However, in case the Fed does not have an information advantage over markets, but signalling its time-varying preferences then purging surprises with respect to the central bank information set would not clean them form information effects.

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 separated into paragraphs and sentences. Topic distributions for the most part correctly capture the meanings of each sentence and paragraph. Moreover, aggregated topic distributions over all the documents are approximately the same as if I were to assign a topic based on the threshold 0.3 for each sentence and 0.25 for each paragraph (see Figure C.29, Figure C.30 and Figure C.31).

Figure 1 shows aggregated topic distributions over all the documents with topics assigned for each sentence. Based on the results, the Fed provides the greatest amount of information on its monetary policy (topic: Fed), economic conditions (topics: Economic, Economics), federal committee regulations (topic: Rules), interest rates setting (topic: Rates), reporting (topic: Reports), job market conditions (topic: Jobs), asset market (topics: Investing, Securities), budget (topics: Income, Taxes, Budget, Spendings), and oil/gas (topics: Gas, Energy, Oil prices, etc.).

These topics are in line with types of information the Fed usually releases in its statements. Infrequent and non-intuitive topics might reflect changes in information that the Fed releases. For instance, the Health topic time series is important when the Fed talks about the effect of Coronavirus in its statements; the Food topic time series highlights periods when the Fed talks about food prices; the Computers topic time series might pin down words about monitor or monitoring in the Fed statements; the Housing topic time series might indicate periods when the Fed talks about house prices, etc.

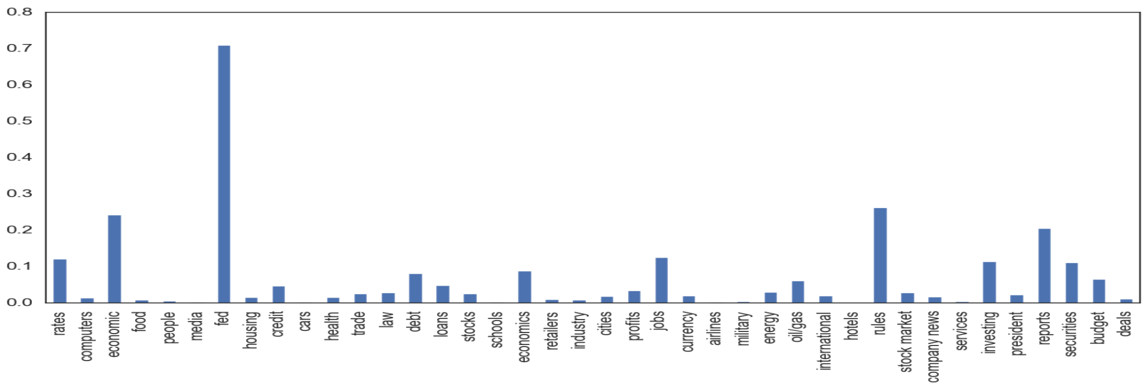
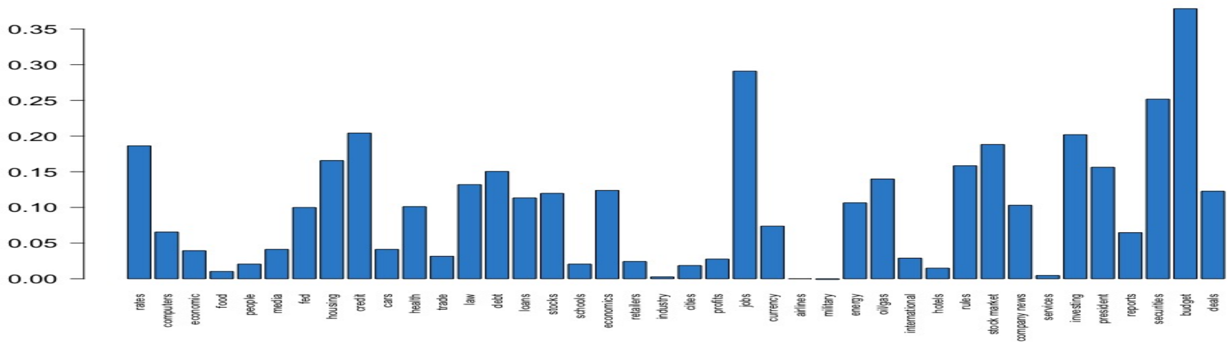


Figure 1: Topic proportions of statements by each sentence

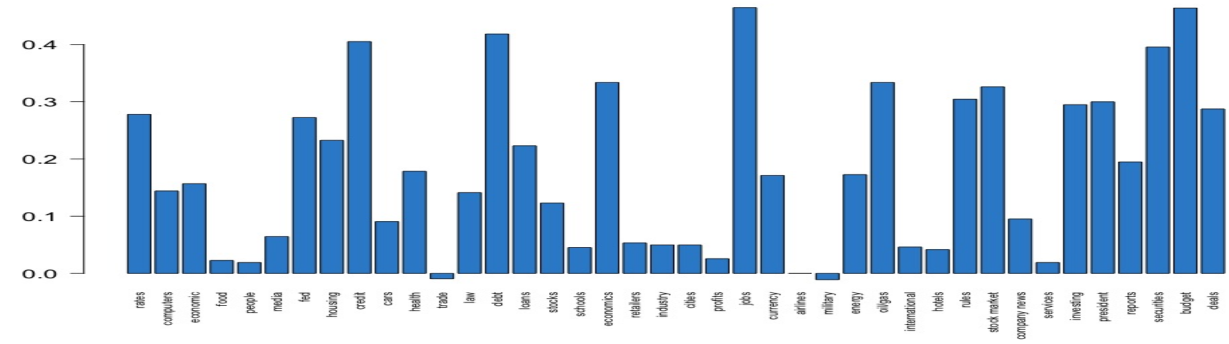
The topic decompositions for FOMC statements over time show that from 2008 the FOMC started to rely more on communications (Figure D.7). That is fully in line with the fact that the federal funds rate hit the zero lower bound and the FOMC started to use unconventional monetary policy tools. The Fed started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

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

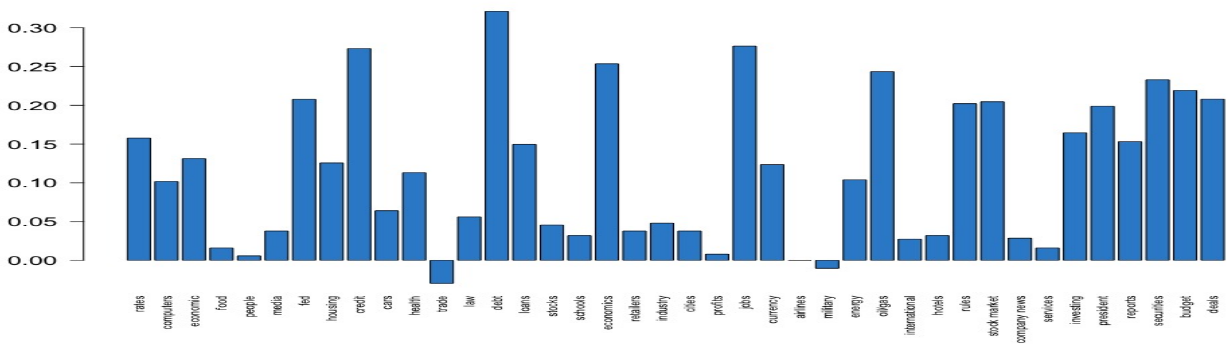
Figure D.8 and Figure D.9 present tone adjusted separate topic time series. Figure 2 reports on changes in topic proportions in Fed statements over time. For instance, the Fed started to signal more regarding jobs, budget, securities, shock market, investing, housing, credit and rates after the funds rate hit the zero lower bound. The rates topic reflects that the Fed started to explain more its interest rate setting decisions, topics concerning budget, securities, shock market, investing, housing, credit might reflect the use unconventional monetary policy tools, while the topic concerning jobs should provide information for the public about future labor market conditions. Therefore, topic time series provide evidence that the Fed started to rely more on communications strategies during unconventional times.



(a) Changes after 2008



(b) QE dates compared to others



(c) QE dates compared to after 2008

Figure 2: Topic proportions over time

The methodology allows to identify topics connected with quantitative easing announcements of the central bank. Figure 2 shows that on dates of these announcements the Fed signaled more on debt, stock market, securities, credit, budget, housing.

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[10]. 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 but not all information is 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 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. It is necessary to set a threshold for selecting the most important topics. In this instance, I use the threshold 0.65, selecting 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, economics, international, company news, investing and deals topics. The result is robust also with regard to important topic time series for the first principal component of the surprises in fed funds futures and eurodollar futures with one year or less to expiration. Importance of topics on economic issues might contain the Fed information effect. For instance Jarocinski & Karadi (2020) found that a difference between the staff and private forecasts for one-quarter-ahead real GDP growth influences the central bank information shocks significantly.

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).

[10]Usually it is a 30-minutes window around the announcement time.

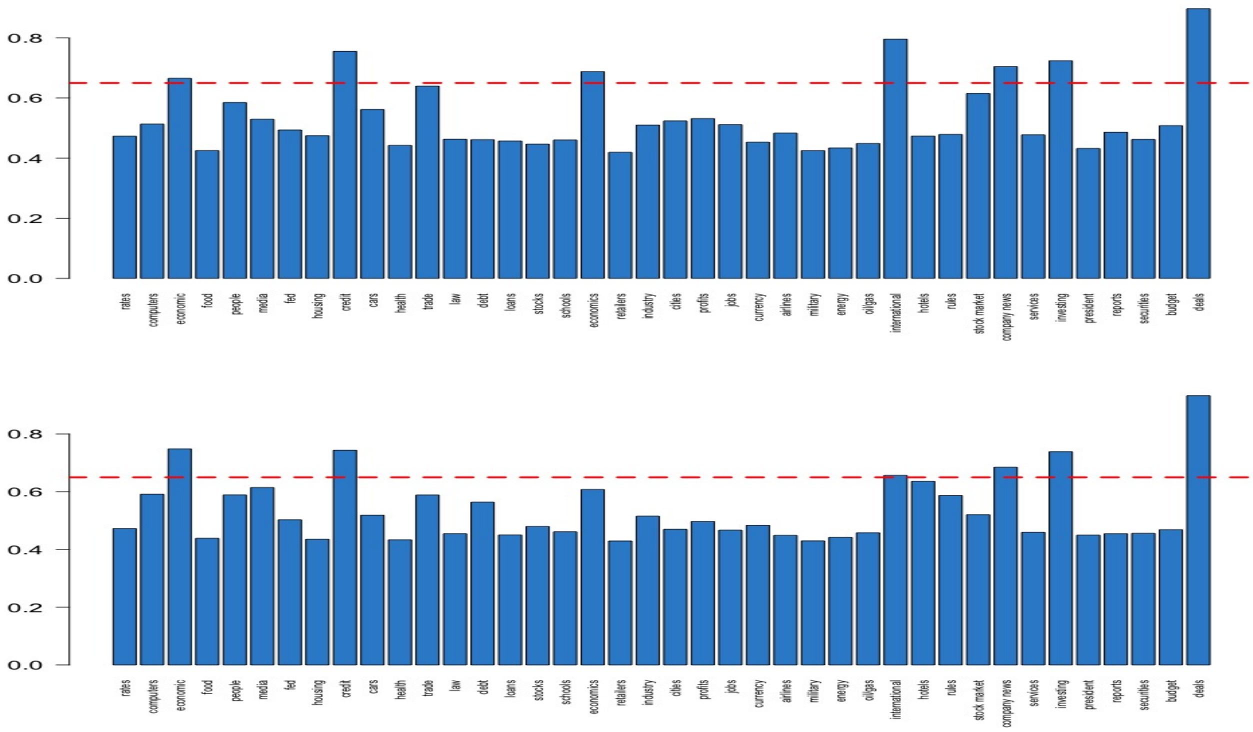


Figure 3: Bayesian Lasso for surprises in 3m federal funds futures (top) and the first principal component of the surprises in fed funds futures and eurodollar futures with one year or less to expiration (bottom)

The Fed does not talk about cities in its statements, rather Cities is purely a label to serve as a topic for the distribution of words. The topic Cities represents sentences that contain a particular combination of words, such as: 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 necessarily need to contain information about cities per se, but there might be information concerning development, projects, etc. This topic is quite infrequent in FOMC statements (Figure 2). Similarly, the topic Cars is not limited exclusively to cars, but also covers 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.

The topics relating to trade and industry which contain uncertainty sentiments are also found to be important for surprises in Federal funds futures (Figure E.1). Furthermore, the topics Computers, Economic, Health, Trade, Industry, Cities, Services, Investing and Deals are found to be important for Gertler & Karadi (2015) proxy for surprises in federal funds futures (Figure E.1).

Figure 4 sheds a light on asymmetric effects of Fed information on surprises in federal funds futures and S&P500 in a narrow window around announcements. Surprises in federal funds futures are more susceptible to negative Fed signals on economy, credit, economics and investing, while surprises in S&P500 are influenced by positive signals concerning the Fed, health, stocks and securities, and by negative signals on credit, trade and currency. In line with logic, surprises in S&P react more on signals about stock markets, whereas surprises in federal funds futures on signals about the economy.

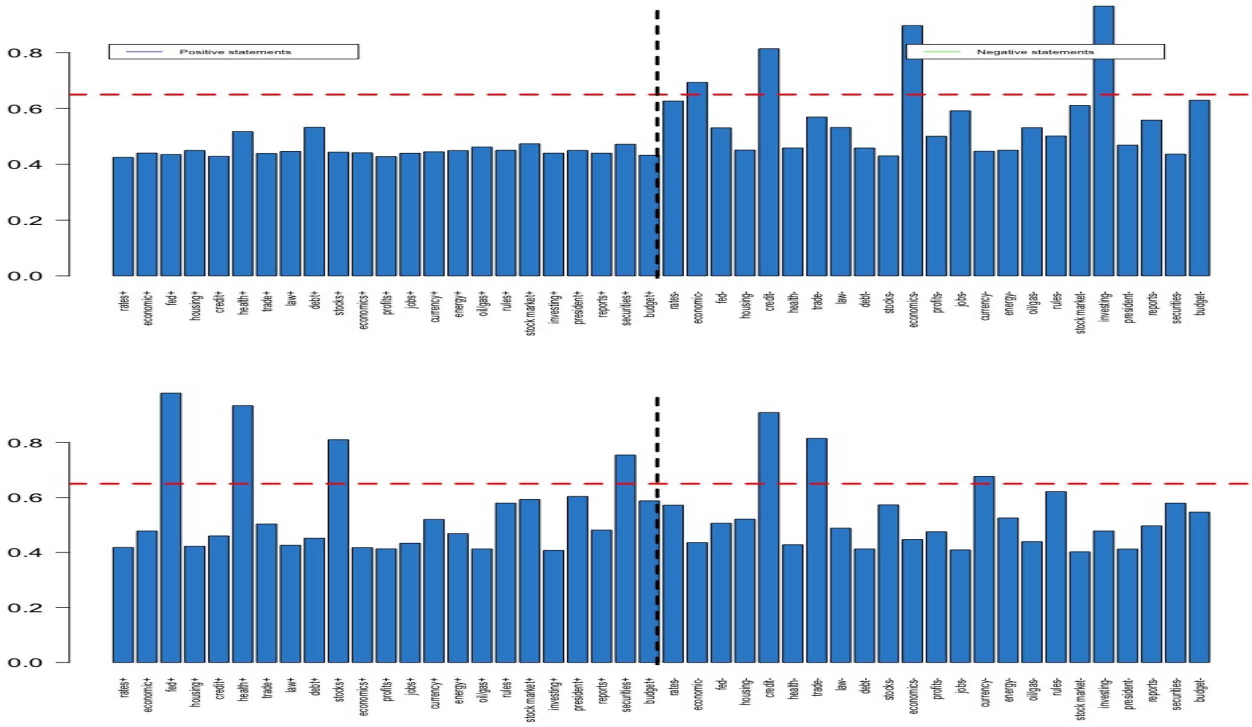


Figure 4: Bayesian Lasso for surprises in 3m federal funds futures (top) and S&P500 (bottom). Asymmetric effects of information

Figure E.2, Figure E.3, Figure E.4, Figure E.5 discuss further asymmetric effects of central bank statements.

4 Mechanism of Central Bank Communication effects

The previous section stated that information released by the central bank in its statements is important for expectations revisions by the public. But what is the channel of propagation of central bank communication on the economy? Central bank communication aims to shape agents expectations of future interest rates and economic conditions, and, therefore, communication should affect through expectation revisions. However, Hansen et al. (2019) showed that news on economic uncertainty can have increasingly large effects along the yield curve. The authors argued that these central bank's signals that drive long-run interest rates do not affect short-run rates and operate primarily through the term premium and have an effect through shaping perceptions of long-run uncertainty.

Firstly, I argue that the central bank sends signals about current macroeconomic conditions and, therefore, it is plausible to take them into account in the standard VAR instead of conditioning on central bank information set, which is unavailable to the public in real-time. I show that central bank communication can be predicted from forward-looking financial market variables. As forward-looking variables, I use changes between FOMC meetings in nominal effective exchange rate (Δ NEER) for USA and Euro, TED Spread (Δ TEDRATE), which is calculated as the spread between 3-Month LIBOR based on U.S. dollars, and Moody's Seasoned Aaa and Baa Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity (Δ AAA10Y and BAA10Y). Series were downloaded from *Federal Reserve Economic Data* (2019).

Table E.1, Table E.2, Table E.3, Table E.4 report results of predictive regressions for all forty topic time series from FOMC statements. I concentrate on topic time series from FOMC statements that are (1) connected to news about the economy, (2) important for surprises in the federal funds futures in a narrow window around announcements, and that are not important for surprises in S&P500 during announcements. The topics are Economic, Economics and Investing. Table E.1 and shows that the Economic topic from statements is predicted from changes in the spread between 3-Month LIBOR based on U.S. dollars and Moody's Seasoned Baa Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity. Both variables serve as indicators of credit risks. The Economics topic from Fed statements can be predicted from changes in S&P500 (Table E.2), which is a stock market index that tracks 500 large companies. According to Table E.4 the Investing topic from FOMC is also predicted from changes in the spread between 3-Month LIBOR based on U.S. dollars. These results are in line with the recent findings of Beckers (2020), who claimed that credit risk conditions enter the central bank reaction function.

Table 1 shows the connection between aggregated signals about the economy[11] in FOMC statements and surprises in 3-month federal fund futures around Fed announcements. Interestingly, R^2 from these regressions are similar to R^2 in the first stage regressions of Miranda-Agrippino & Ricco (forthcoming), who regressed surprises around FOMC announcements on Fed private forecasts.

Table 1: Surprises in ffr futures

	Dependent variable:	
	ffr_hf	ffr_hf_PCA
	(1)	(2)
Economic aggregated	1.696**	2.174**
	(0.831)	(0.940)
Constant	−0.002	0.009***
	(0.002)	(0.003)
Observations	274	274
R^2	0.098	0.086
Adjusted R^2	0.095	0.083
Note:	Newey-West HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01	

Furthermore, it is important to investigate whether Fed signals are important for interest rate changes. For this, I aggregate the topics Economy, Economic and Investing into one and see if it has predictive power for interest rate changes around Fed announcements days[12]. Table 2 reports on the importance of economic news signals sent by the Fed during announcement days on daily changes in short-term rates. As a baseline, I use two-days difference in short-term rates, i.e. one day after an announcement minus one day before an announcement. That is because, as noted in the literature, markets might need time to adjust for the information beyond the Fed action itself.

[11] I aggregate the topics Economy, Economic and Investing into one

[12] Daily yields are taken from Gürkaynak et al. (2007)

Central bank's signals about the economy in its statements are positively correlated with changes in short-term yields. This might indicate the expectations channel of central bank communication. Positive signals of central bank concerning the economy lead to a revision of expectations by market participants. If central bank information set was the same as market participants information these signals would not be important for markets. Moreover, these signals can be predicted by forward-looking financial variables and according to the efficient market hypothesis should already be taken into account by markets by release date.

Table 2: Δ Yields, 2 days difference

	Dependent variable:				
	1 year rate Δ 1 day	1 year rate	breakeven 5 years	breakeven forward 2 years	breakeven 10 years
	(1)	(2)	(3)	(4)	(5)
Economic aggregated	1.294*	1.733**	−1.563	−2.260*	−1.046
	(0.757)	(0.782)	(1.134)	(1.311)	(0.684)
Constant	0.0004	−0.008	0.005	0.012*	0.001
	(0.004)	(0.005)	(0.005)	(0.007)	(0.004)
Observations	186	186	132	132	132
R^2	0.039	0.048	0.036	0.054	0.029
Adjusted R^2	0.034	0.042	0.028	0.046	0.022

Note: Newey-West HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Reverse causality is not valid in this case, because the Fed should not react to previous day yields in its statements. Moreover, the assumption of no omitted variable bias is not too restrictive because usually there are no other events during the days of statements releases. One concern is that markets react to unexpected actions of the central bank and these actions are correlated with signals concerning the economy. But the more plausible explanation, in this case, is that markets react to signals and to actions at the same time. The latter claim was confirmed by Gürkaynak et al. (2005) for instance.

Moreover, as was notices by Hansen et al. (2019), central bank signals are highly-dimensional objects, which can affect the term premium as well. The authors pointed out the importance of central bank signals concerning macroeconomic uncertainty. To study this channel I use topic time series that were labelled with uncertainty sentiments instead of positivity/negativity. Moreover, I concentrate on topics concerning the economy, because these topics should reflect central bank views about future macroeconomic conditions and, therefore, these should be a source of important information for markets.

Appendix F reports on the predictive power of central bank uncertain signals concerning the economy for the yield curve, Treasury Inflation-Protected Securities (TIPS) and breakeven inflation rates[13]. Economic aggregated is the sum over uncertainties from the Economic, Economics and Investing topics. I also control for uncertainties from the Fed topic that should capture Odyssean forward guidance, where the Federal Reserve release uncertainty concerning information about future monetary policy. Fed topic on FG dates controls for this topic on days of forward guidance.

The results show that uncertainty concerning the economy released by the Fed in its statements affects the long tail of the yield curve. That is in line with the finding of Gürkaynak et al. (2020), who showed that a statement is more informative for longer maturities. Uncertainty concerning the economy is positively connected with daily changes in ten, fifteen, twenty, twenty-five and thirty years yields. The result is robust also for two days changes in yields and also while controlling for surprises in federal funds futures. The results are completely in line with those of Hansen et al. (2019), who found that long-run interest rates respond to central bank communication, namely central bank uncertainty signals on economic development. The authors used the publication of the Bank of England's Inflation Report. The uncertainty signals that drive long-run interest rates do not affect short-run rates and operate primarily through the term premium.

Central bank uncertain signals concerning the economy are also positively connected with two five and ten years forward rates and with one-year forward rate four years ahead. The result is also robust to controlling for a measure of surprises in federal funds futures and to different ways of differencing dependent variables. Moreover, central bank uncertain signals concerning the economy are positively connected to all curve of Treasury Inflation-Protected Securities, while surprises in federal funds futures on announcement days are not. Therefore, uncertain signals concerning the economy released by the Fed in its statements might affect the yield curve of real interest rates.

[13]Daily yields, TIPS and breakeven inflation rates are taken from Gürkaynak et al. (2007) and Gürkaynak et al. (2010). Inflation compensation incorporate inflation risk premiums and the effects of the differential liquidity of TIPS and nominal securities.

Appendix F.2 presents the robustness check results where I include also a measure of surprises in a narrow window around announcements[14]. All results concerning the importance central bank uncertain signals hold. Surprisingly enough, the measure of surprises about future interest rates is neither connected with changes in yields of Treasury Inflation-Protected Securities, nor with daily changes in Inflation Compensation.

Appendix F.3 discusses bad controls and measurement errors issues. Measurement error is a potential issue with the results because the coefficient of surprises in federal funds futures is higher when I add my proxies for central bank uncertainty signals compared to coefficients from univariate regressions. In this case, there might be a slight upward bias. Bad controls situation occurs when potential outcome variables are used as controls in a regression. I show that surprises in federal funds futures around announcements cannot be outcome variables in a regression.

Hanson & Stein (2015) argue that news about short-term policy expectations is propagated to longer-maturity bonds by the trading activity of yield-oriented investors. According to their model, decreases in short rates induce these investors to switch to longer-maturity bonds, driving the yields on such bonds down through changes in the term premium. Hansen et al. (2019) found that central bank communication affects long-run interest rates by providing news on risk and uncertainty around economic conditions, and thereby generating a change in the long-run term premium. This channel operates not by changing long-run expectations of economic conditions, but by changing the perceived variance of those conditions. Furthermore, the effect of uncertainty signals comes via the long-run term premium, which can move independently of short-run expectations. My results confirm those of Hansen et al. (2019), central bank communication indeed affects market beliefs about long-run uncertainty.

[14]Here 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 because it should capture more of forward guidance.

5 Monetary policy vs. Information shocks

5.1 Baseline results

Following Jarocinski & Karadi (2020), I use Cholesky identification[15] for monetary policy shocks with Jarocinski & Karadi (2020) original variables in the following order: surprises in 3-month federal funds futures, the one-year government bond yield, real GDP, GDP deflator and the excess bond premium. To separate a pure monetary policy shock from an information shock, I add additional information variables before surprises in 3-month federal funds futures. The studied period is 1994:M3–2016:M12. Because the data are at monthly frequency I use twelve lags in SVAR. Appendix G presents the SVAR estimation details.

As information variables I select those that should capture the effects of information about the economy, that are topics concerning economy, economics and investing. These topics time series were selected based on following criteria: (1) they have high predictive power for surprises in federal funds futures; (2) they are connected to news about the economy as opposed to monetary policy decisions per se; (3) they are not connected to quantitative easing announcements, which are mainly concerning debt, housing, stock market and securities. Moreover, these topic time series affect the yield curve during the announcement dates.

Figure 5 discusses the baseline results, focusing on three distinct types of shock. In Panel (a) the surprises in 3-month Federal funds futures are ordered first; in Panel (b) the information variables are ordered before the surprises in 3-month Federal funds futures; while Panel (c) presents the difference between the two, which should capture information effects.

The baseline results (Panel (a)) are fully in line with the results of Jarocinski & Karadi (2020) and Gertler & Karadi (2015). Some difference in magnitudes might be explained by their use of a different period of study as Jarocinski & Karadi (2020) used the period from 1984 and employed Kalman filter and smoother for substituting the missing values in surprises in 3-month Federal funds futures. Also, prior tightness parameters are a bit different since I use tighter prior for lags further than the first one.

[15]The authors use Cholesky identification as alternative specifications to sign restrictions.

The result of a small decline in S&P500 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 lesser amount in response to the FOMC announcement than to the shock without information about future exogenous fundamentals.

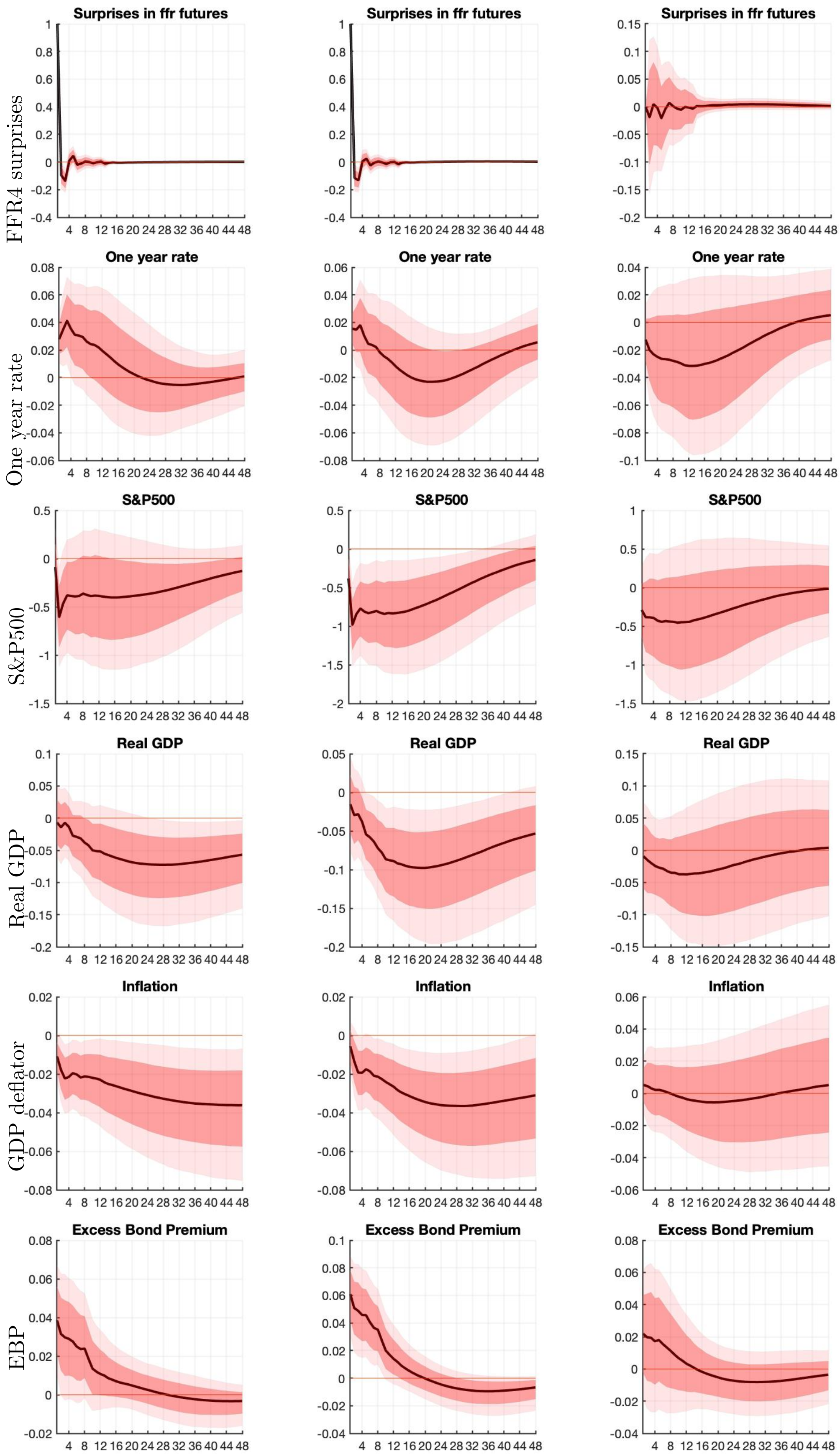
Panel (b) presents the results for purged shocks, which should not contain the Fed information effect. The results are similar to Jarocinski & Karadi (2020). The response of the one year rate is more transitory than in Panel (a). The response of S&P500 is negative for the first few months. The response 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) with the consequence that the identified effect looks like a contractionary monetary policy shock. For all these variables there is a higher posterior probability for a contractionary response because even 90% posterior credible sets are below zero for a long period. The response of the Excess Bond premium is also in line with the results Jarocinski & Karadi (2020) - a contractionary monetary policy shock without an information effect has a greater effect on the cost of credit with narrower credible sets, which lasts for almost twelve months.

Panel (c) discusses the results for differences between two previous effects, which should capture pure negative information shocks. The results are in line with those of Jarocinski & Karadi (2020) and Steinsson (2019): the Fed information shock has a more prolonged but muted effect on the one year rate, on S&P500 and real GDP. An interesting result is that it has a large short-run effect on the EBP. So positive news contained in announcements can reduce the costs of credit in the short-run.

Therefore, I studied the effect of information free policy shock without relying on sign restrictions. Some differences in impulse responses from Jarocinski & Karadi (2020) ones might be explained by (1) differences in the periods studied[16], (2) and the identification strategies for monetary policy and information shocks[17].

[16]Jarocinski & Karadi (2020) dealt with missing values for the shocks series via Kalman filter and smoother.

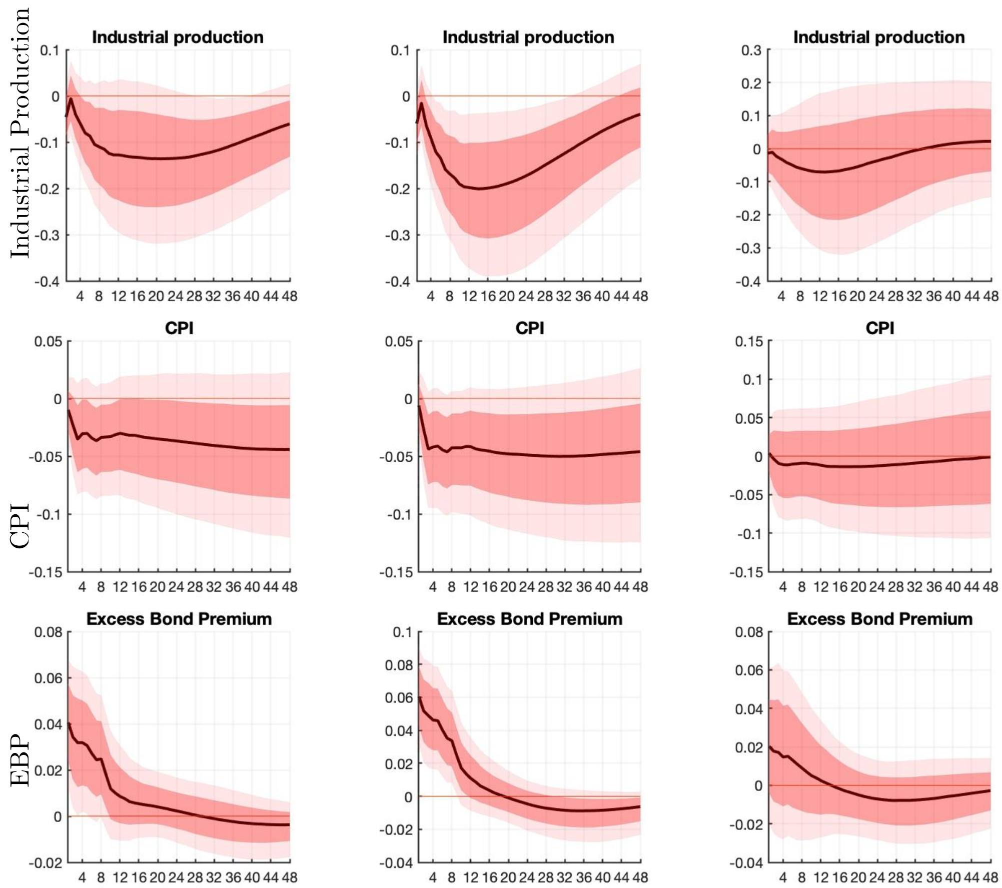
[17]Jarocinski & Karadi (2020) employed sign restriction which is set identification while Cholesky is point identification.



(a) Baseline (b) Policy shock (c) Difference

Figure 5: Comparison between monetary policy and information shocks. 3m federal funds futures shaded 5%,16%, 84% and 95% percentiles

Figure 6 presents the robustness exercise with the log U.S. consumer price index and industrial production as proxies for inflation and economic activity. The results are completely in line with the baseline results in Figure 5, namely industrial production and inflation decline in response to a policy shock with higher posterior probability in case of controlling for informational effects. The difference between information free policy shock and a policy shock is visible the most in the Excess Bond premium, but differences in responses of real economic activity are also distinguishable.

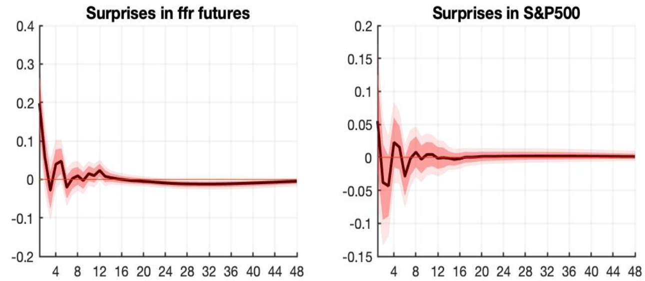


(a) Baseline (b) Policy shock (c) Difference

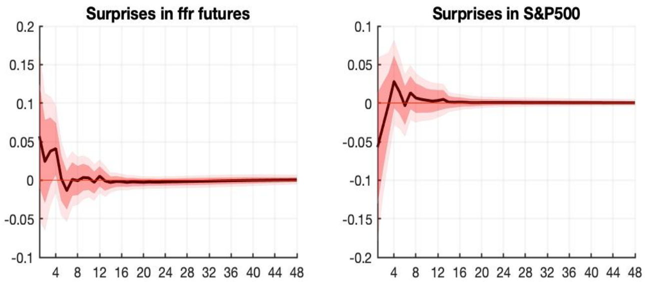
Figure 6: Comparison between monetary policy and information shocks. 3m federal funds futures shaded 5%,16%, 84% and 95% percentiles

Appendix H discusses the results of Forecast error variance decompositions to both shocks and their difference. A pure monetary policy shock explains higher proportion of the forecast error variance of the Excess Bond Premium during the whole period, higher proportion of S&P500 just after the shock and higher proportion of GDP and industrial production in the long-run. It also explains a lower share of the one-year rate on impact.

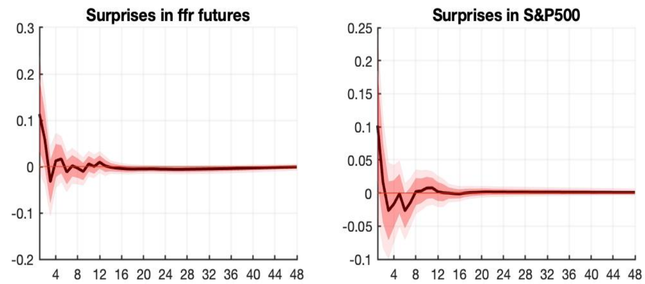
Figure 7 shows the responses in surprises in 3-m federal funds futures and S&P500 to information shocks. According to the theory, they should respond to information shocks in the same direction. It is seen that mainly high-frequency surprises respond to a positive economic information shock as the theory predicts. This economic information might capture the effects of Delphic forward guidance.



(a) Information shock 1



(b) Information shock 2

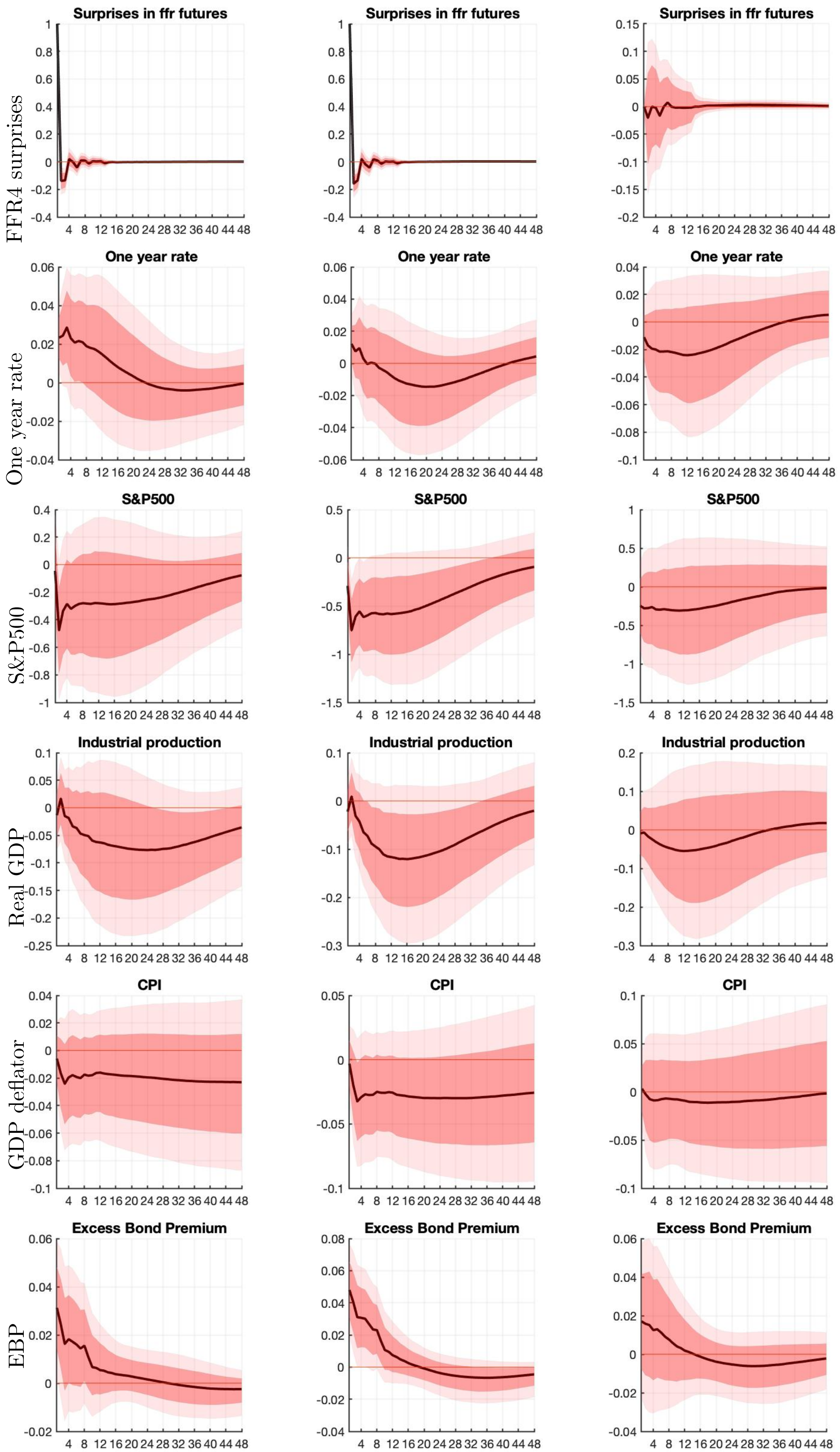


(c) Information shock 3

Figure 7: Information effects

5.2 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)).



(a) Baseline (b) Policy shock (c) Information shock

Figure 8: 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

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 8 presents the results.

There are differences between these results and the baseline results from Figure 5. For responses of industrial production and CPI there is a posterior probability mass that lies within a region of positive values while looking at 68% credible sets. For a response of the excess bond premium, there is a region of 90% posterior probability mass that takes negative values. Adding information variables to VAR reduces these probabilities of incorrect impulse responses and sharpens identification.

The results concerning the effect of information-free shock are similar to the results from Figure 5, with the exception of 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 the information shock having a larger effect 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.

5.3 Transmission of monetary policy shocks

To study the transmission of monetary policy shocks I use large-scale Bayesian VAR following the work of Banbura et al. (2010), who introduced dummy variables prior similar to Minnesota to work with a large number of variables in VAR. The model assumes natural conjugate Normal-Inverse-Wishart prior for autoregression coefficients and variances. As hyper-parameters I use λ 0.1[18], that controls overall prior tightness, and as the prior means of coefficients I use ones for trending variables and zeros for stationary variables, prior mean for a constant is 100[19].

Figure 9 presents the results for a medium-scale VAR, that includes information variables, monetary policy shock variable, various interest rates and expectations from Consen-

[18] I tuned this hyper-parameter to match impulse responses of small scale VAR with Independent Normal-Inverse-Wishart prior with Minnesota hyper-parameters, as shown in Appendix G

[19] These are the conventional settlings in literature

sus Economics. Surprisingly, even without short-run restrictions slow-moving variables do not respond much on impact, while fast-moving variables respond more sharply on impact. The shock increases the costs of credit for about four months after the shock and reduces S&P500 for about eight months after the shock. It leads to a steady decline in inflation and a decline with reversion in real economic activity. The shock also leads to more negative expectations of GDP and inflation. As a result, longer-term interest rates are not rising to reflect these expectations. Moreover, it does not seem that a contractionary shock transmits through term premium.

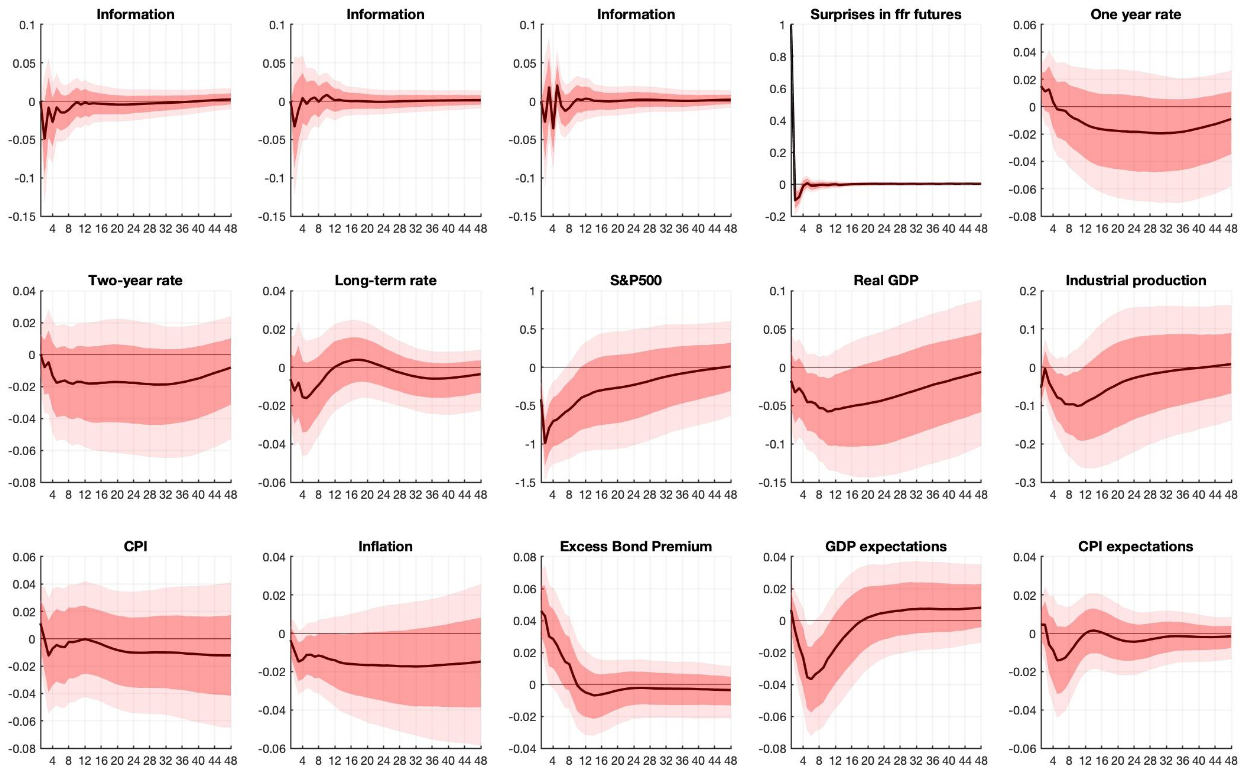


Figure 9: Monetary policy shock in medium-scale VAR

I use the big database of McCracken & Ng (2015), that contains 128 monthly variables. This database includes variables that should capture variables from the central bank's reaction function. Moreover, there are forward-looking financial variables, which should capture central banks' and agents' foresight. All variables were transformed into stationary form, following the recommendations of McCracken & Ng (2015), and afterwards impulse responses were cumulated in levels.[20] Figure 10, Figure 11 and Figure 12 report the results.

All slow-moving variables do not respond much on impact without any zero contemporaneous restrictions, while fast-moving variables do respond on impact. The shock leads to an increase in all short-term interest rates up to one year rate.

The findings confirm the interest rate channel of monetary policy transmission: industrial production of durable consumer goods, business equipment and durable materials fall more compared to other components of industrial production. Moreover, employment in the durable sector falls at a higher rate than corresponding employment in the non-durable sector. New orders for durable goods rise on impact after the shock, while at the same time unfulfilled orders for durable goods increase for a few months after the shock. Inflation in the durable sector does not decline in response to a shock, while the growth rate of personal consumption expenditures on durables declines on impact.

In line with the results of Gertler & Karadi (2015), the credit channel[21] is found to be an important channel of monetary policy propagation. The shock rises on impact the Excess Bond Premium, the three-month commercial paper spread, and leads to an increase in the long-run spread between Moody's BAA and the effective federal funds rate. The Excess Bond Premium reflects long-run borrowing costs in the non-farm business sector, the three-month commercial paper spread is relevant to the cost of short-term business credit and the cost of financing consumer durables, and BAA spread measures credit risk. The shock leads to a decrease in commercial loans on impact.

The balance sheet channel cannot be estimated directly. Nevertheless, the shock leads with high probability to a reduction in house prices, since the largest share of posterior probability mass lies in the negative region. That might be explained by the fact that higher interest rates increase the costs of owning a house, which implies a lower asset value. Therefore, a lower value of collateral leads to rising the borrowing cost, making it harder for smaller or younger firms to get access to credit through asymmetric information among economic agents.

As for the expectation channel of monetary policy, a contractionary monetary policy shock also leads to a decline in consumer confidence in the long-run, measured by the

[20]With the exception of variables that were double differenced. Impulse responses for these variables are in growth rates.

[21]Bernanke & Gertler (1995)

consumer sentiment index. The shock also increases macroeconomic uncertainty, but the effect is not persistent and disappears in about twelve months after the shock.

The asset price channel is also visible from impulse responses: S&P500 declined in the long-run in response to a contractionary monetary policy shock, as well as S&P industrial. Lower asset prices together with lower in house prices, lead to a decline in consumption and investment via wealth effect and the effect on the value of collateral. A decline in manufacturing capacity utilisation could lead to subdued business investment in the future. Moreover, there is a negative growth rate of personal consumption expenditures on nondurable goods in about four months after the shock.

U.S. dollar appreciates on impact based on the response of trade-weighted U.S. dollar index, that is compared to weighed shares of Euro, Japanese yen, Canadian dollar, British pound, Swedish krona, and Swiss franc. That confirms the importance of the exchange rate channel of monetary policy.

According to the results, all components of industrial production steadily decline in response to a contractionary shock. Capacity utilisation in manufacturing also steadily falls meaning that actual output in manufacturing slowly falls with respect to its potential level.

Unemployment starts to increase in about eight months after the shock together with average unemployment duration. The impact on unemployment is not distinguishable over the short-term, possibly due to nominal rigidities in the economy. This effect mainly leads to a larger share of long-term unemployed people. Average weekly hours worked also start to decline in about eight months after the shock but the effect is less persistent here. Initial claims increase in about four months after the shock and the effect is persistent for about four years. These claims are filed by an unemployed individual after a separation from an employer for eligibility for the Unemployment Insurance program.

A monetary contraction causes all components of growth rates of inflation to decline in the short-run with exception of apparel, medical care and durables. These components are less sensitive to a monetary policy shock.

Total business inventories start to decline steadily a few months after the shock, but sales decline at a higher rate and therefore total business inventory to sales ratio increases from the fourth to the twelfth months after the shock.

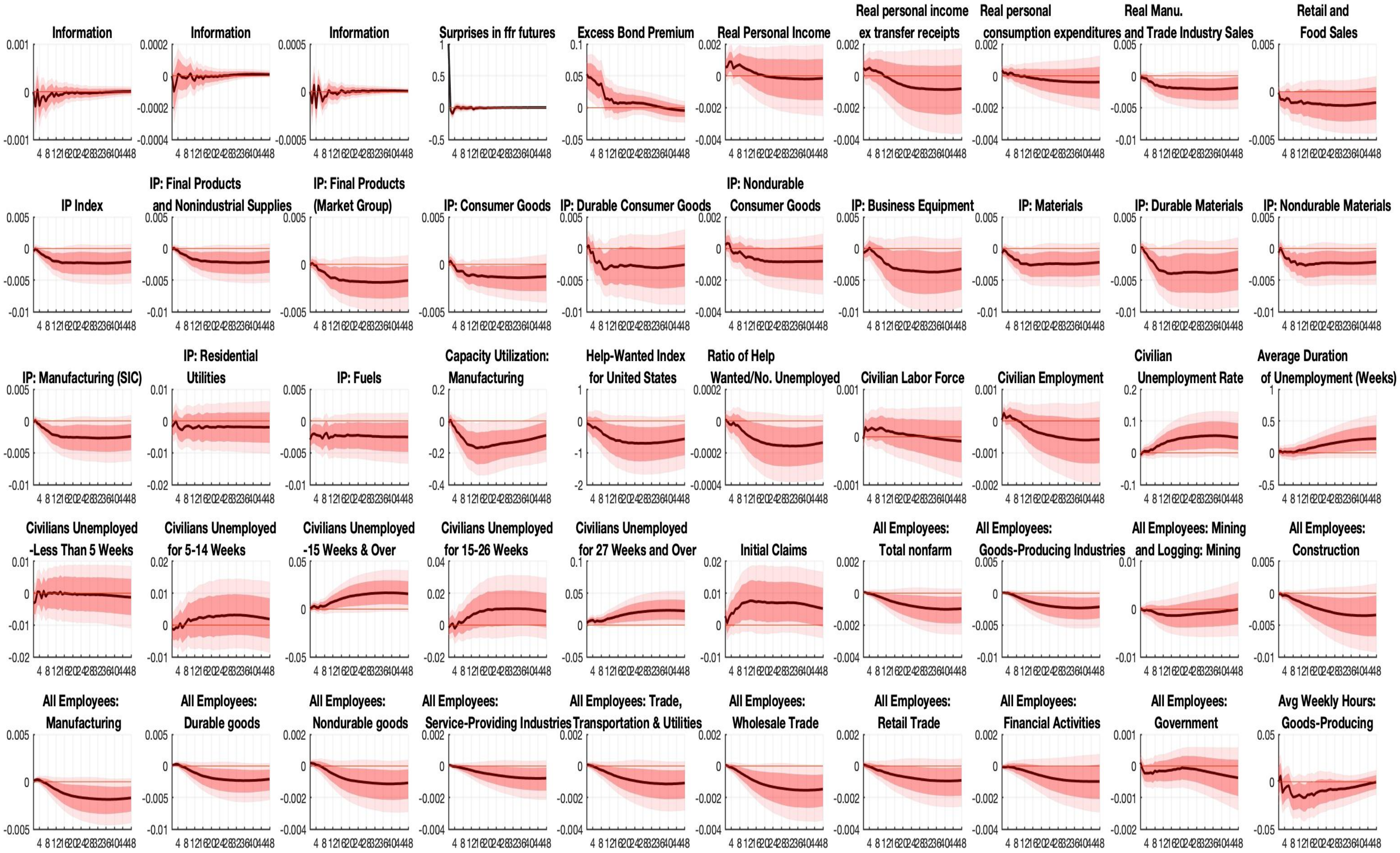


Figure 10: Monetary policy shock in large-scale VAR

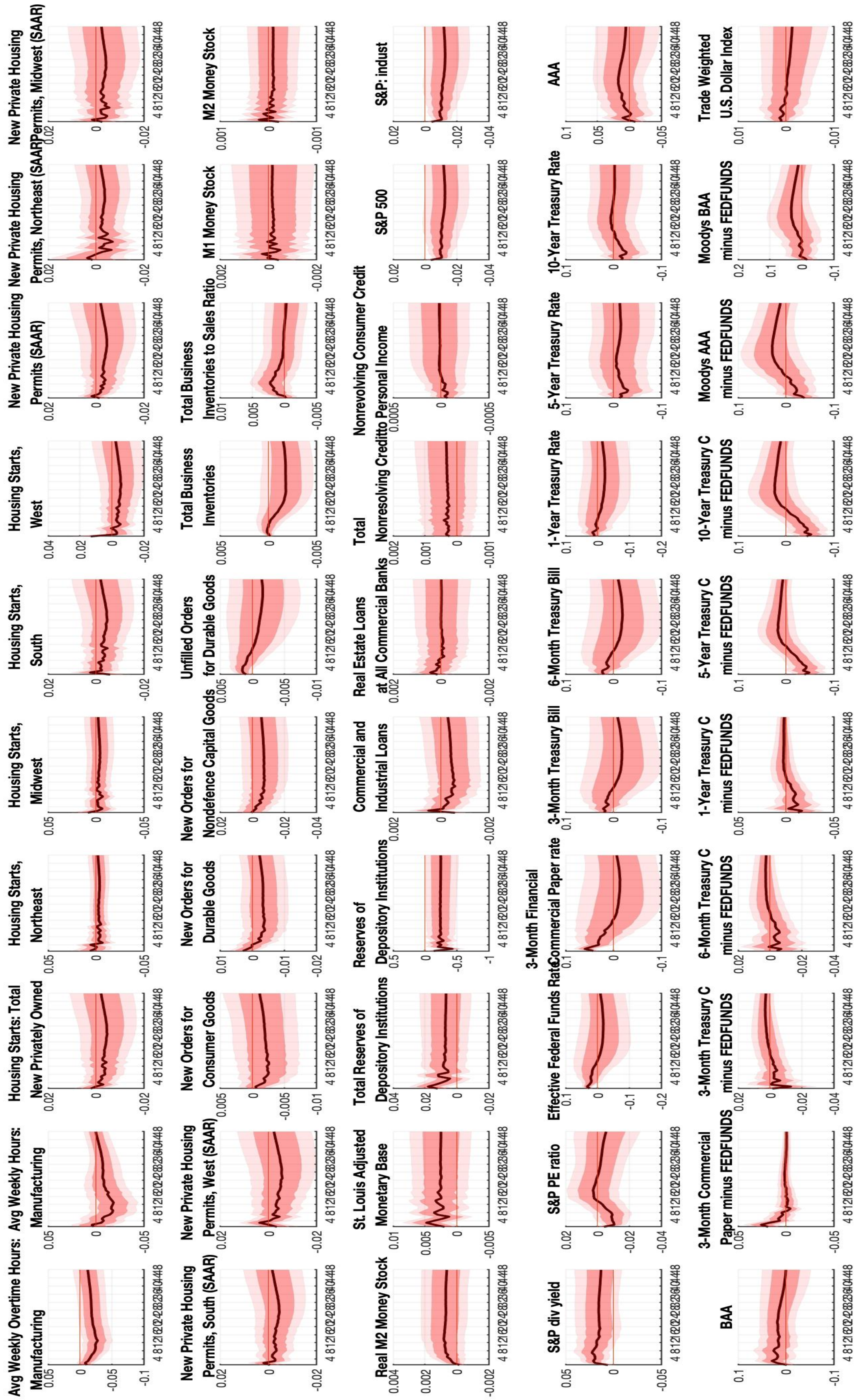


Figure 11: Monetary policy shock in large-scale VAR

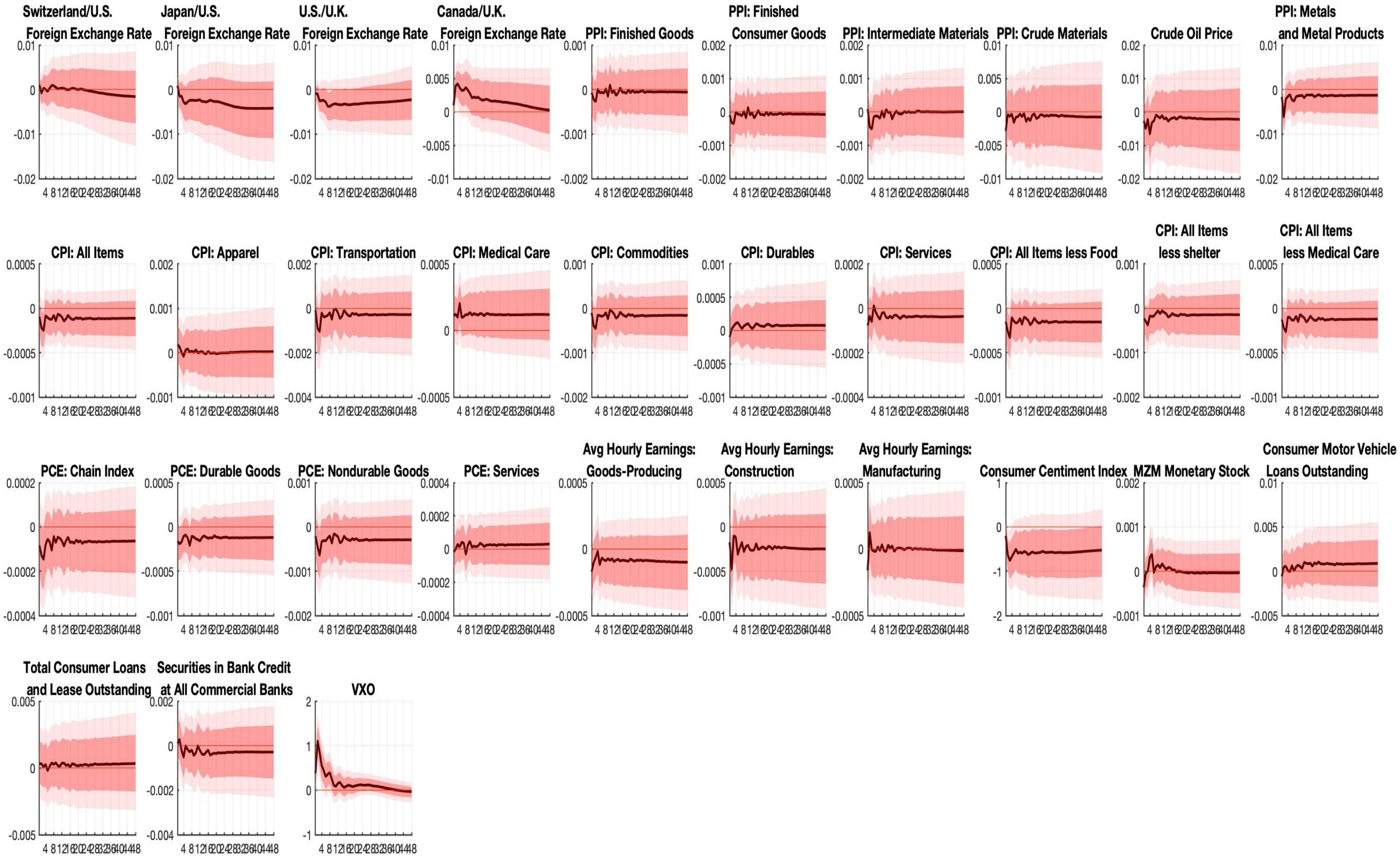


Figure 12: Monetary policy shock in large-scale VAR

6 Conclusions

The paper elaborates on the recent contribution of Jarocinski & Karadi (2020) in decomposing information from policy shocks, as well as on invertibility problem when econometrician's information set differs from decision maker's. This study uses information from FOMC statements and augments the standard VAR with important information. This allows to condition on information that is directly available to the public in real-time.

The study combines topic time series from FOMC statements with the tone of these statements. I extract information from FOMC statements by using Latent Dirichlet Allocation that was pre-trained on the business section from major U.S. newspapers. The 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.

I use information released by the Fed in its' statements as additional variables in VAR that might affect policy surprises contemporaneously. The results show that a policy shock has a more negative effect on GDP, a more prolonged negative effect on inflation and greater effect on the excess bond premium compared to the baseline surprises measure. In the short-run it causes S&P500 to decline and the Fed to raise its interest rate. The transmission channels of monetary policy identified in this paper are in line with the theory: monetary policy operates through the interest rate, credit, asset prices, exchange rate and expectations channels. What is more, I did not find evidence of the importance of the term premium channel for monetary policy transmission.

References

Arthur, D. & Vassilvitskii, S. (2007), 'k-means++: The Advantages of Careful Seeding', *SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms* pp. 1027–1035.

Bahmani, B., Moseley, B., Vattani, A., Kumar, R. & Vassilvitskii, S. (2012), 'Scalable K-Means++', *Proceedings of the VLDB Endowment* **5**(7), 622–633.

Banbura, M., Giannone, D. & Reichlin, L. (2010), 'Large Bayesian vector auto regressions', *Journal of Applied Econometrics* **25**(1), 71–92.

Barsky, R. B. & Sims, E. R. (2012), 'Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence', *The American Economic Review* **102**(4), 1343–1377.

Beaudry, P. & Portier, F. (2006), 'Stock Prices, News, and Economic Fluctuations', *The American Economic Review* **96**(4), 1293–1307.

Beaudry, P. & Portier, F. (2014), 'News-driven business cycles: Insights and challenges', *Journal of Econometric Literature* **52**(4), 993–1074.

Beckers, B. (2020), 'Credit Spreads, Monetary Policy and the Price Puzzle', *Research Discussion Paper Economic Research Department Reserve Bank of Australia* (2020-01).

Beechey, M. J. & Wright, J. H. (2009), 'The High-Frequency Impact of News on Long-Term Yields and Forward Rates: Is It Real?', *Journal of Monetary Economics* **56**(4), 535–544.

Benhabib, J. & Spiegel, M. M. (2019), 'Sentiments and Economic Activity: Evidence from US States', *The Economic Journal* **129**(618), 715–733.

Bernanke, B. S., Boivin, J. & Eliasz, P. (2005), 'Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach', *The Quarterly Journal of Economics* **120**(1), 387–422.

Bernanke, B. S. & Gertler, M. (1995), 'Inside the Black Box: The Credit Channel of Monetary Policy Transmission', *Journal of Economic Perspectives* **9**(4), 27–48.

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003), 'Latent Dirichlet Allocation', *Journal of Machine Learning Research* **3**, 993–1022.

Branch, W. A. (2004), 'The Theory of Rationally Heterogeneous Expectations: Evidence from Survey data on inflation expectations', *The Economic Journal* **114**(497), 592–621.

Bu, C., Rogers, J. & Wu, W. (2020), 'Forward-Looking Monetary Policy and the Transmission of Conventional Monetary Policy Shocks', *Finance and Economics Discussion Series 2020-014. Washington: Board of Governors of the Federal Reserve System* (2020-014).

Campbell, J., Evans, C., Fisher, J. & Justiniano, A. (2012), 'Macroeconomic Effects of Federal Reserve Forward Guidance', *Brookings Papers on Economic Activity* **43**(1), 1–80.

Carroll, C. D. (2001), 'The Epidemiology of Macroeconomic Expectations', *NBER Working Paper* (No. 8695).

Carroll, C. D. (2003), 'Macroeconomic Expectations of Households and Professional Forecasters', *Quarterly Journal of Economic* **118**(1), 269–298.

Coibion, O., Gorodnichenko, Y., Kumar, S. & Pedemante, M. (2020), 'inflation expectations as a policy tool?', *Journal of International Economics* (Forthcoming).

Cole, W. M. (2010), 'No News Is Good News: Human Rights Coverage in the American Print Media, 1980-2000', *Journal of Human Rights* **9**(3), 303–325.

D'Amico, S. & King, T. B. (2017), 'What Does Anticipated Monetary Policy Do?', *Federal Reserve Bank of Chicago* (WP 2015-10).

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.

Doms, M. & Morin, N. (2004), 'Consumer Sentiment, the Economy, and the News Media', *Finance and Economics Discussion Series Washington: Board of Governors of the Federal Reserve System* (2004-51).

Drager, L. & Lamla, M. J. (2017), 'Imperfect Information and Consumer Inflation Expectations: Evidence from Microdata', *Oxford Bulletin of Economics and Statistics* **79**(6), 933–968.

Durante, E., Ferrando, A. & Vermeulen, P. (2020), 'Monetary policy, investment and firm heterogeneity', *ECB Working Paper Series* (2390).

Dvorkin, M. A. & Shell, H. G. (2016), 'The Recent Evolution of U.S. Local Labor Markets', *Economic Synopses* (15).
URL: *https://research.stlouisfed.org/publications/economic-synopses/2016/08/01/the-recent-evolution-of-u-s-local-labor-markets/*

Federal Open Market Committee: Transcripts and other historical materials (2019).
URL: *https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm*

Federal Reserve Economic Data (2019).
URL: *https://fred.stlouisfed.org*

Feve, P. & Guay, A. (2016), 'Sentiments in SVARs', *Working Paper Tolouse School of Economics* (TSE 656).

Flack, E., Hoffman, M. & Hurtgen, P. (2019), 'Disagreement about inflation expectations and monetary policy transmission', *Journal of Monetary Economics* (Forthcoming).

Gertler, M. & Karadi, P. (2015), 'Monetary Policy Surprises, Credit Costs, and Economic Activity', *American Economic Journal: Macroeconomics* **7**(1), 44–76.

Gilbert, T., Scotti, C., Strasser, G. & Vega, C. (2016), 'Is the intrinsic value of macroeconomic news announcements related to their asset price impact?', *European Central Bank Working paper series* (1882).

Gilchrist, S. & Zakrajsek, E. (2012), 'Credit Spreads and Business Cycle Fluctuations', *The American Economic Review* **102**(4), 1692–1720.

Gomes, S., Iskrev, N. & Mendicino, C. (2017), 'Monetary policy shocks: We got news!', *Journal of Economic Dynamics and Control* **74**, 108–128.

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.

Gürkaynak, R. S., Kısacikoglu, B. & Wright, J. H. (2020), 'Missing Events in Event Studies: Identifying the Effects of Partially Measured News Surprises', *The American Economic Review* **110**(12), 3871–3912.

Gürkaynak, R. S., Sack, B. & Swanson, E. T. (2005), 'Do actions speak louder than words? the response of asset prices to monetary policy actions and statements', *International Journal of Central Banking* **1**(1), 55–93.

Gürkaynak, R. S., Sack, B. & Wright, J. H. (2007), 'The U.S. Treasury yield curve: 1961 to the present', *Journal of Monetary Economics* **54**(8), 2291–2304.

Gürkaynak, R. S., Sack, B. & Wright, J. H. (2010), 'The TIPS Yield Curve and Inflation Compensation', *American Economic Journal: Macroeconomics* **2**(1), 70–92.

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.

Hanson, S. G. & Stein, J. C. (2015), 'Monetary policy and long-term real rates', *Journal of Financial Economics* **115**(3), 429–448.

Henry, E. (2008), 'Are Investors Influenced By How Earnings Press Releases Are Written', *The Journal of Business Communication* **45**(4), 363–407.

Hirose, Y. & Kurozumi, T. (2012), 'Identifying News Shocks with Forecast Data', *CAMA Working Papers* (2012-01).

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.

Lamla, M. J. & Lein, S. M. (2014), 'The role of media for consumers' inftation expectation formation', *Journal of Economic Behavior and Organization* **106**, 62–77.

Larsen, V. H. & Thorsrud, L. A. (2019*a*), 'Business Cycle Narratives', *CESifo Working Paper* (7468).

Larsen, V. H., Thorsrud, L. & Zhulanova, J. (2020), 'News-driven inflation expectations and information rigidities', *Journal of Monetary Economics* (Forthcoming).

Larsen, V. & Thorsrud, L. (2019*b*), 'The Value of News for economic developments', *Journal of Econometrics* **210**(1), 203–218.

Le, Q. & Mikolov, T. (2014), 'Distributed Representations of Sentences and Documents', *Proceedings of the 31st International Conference on Machine Learning* **32**(2), 1188–1196.

Leduc, S. & Sill, K. (2013), 'Expectations and Economic Fluctuations: An Analysis Using Survey Data', *Review of economics and statistics* **95**(4), 1352–1367.

Leeper, E., Walker, T. & Yang, S.-C. (2013), 'Fiscal foresight and information flows', *Econometrica* **81**(3), 1115–1145.

Lei, C., Lu, Z. & Zhang, C. (2015), 'New on inflation and the epidemiology of inflation expectations in China', *Economic Systems* **39**(4), 644–653.

Levchenko, A. A. & Pandalai-Nayar, N. (2015), 'TFP, News, and "Sentiments": The International Transmission of Business Cycles', *NBER Working Paper* (21010).

Lewis, D. J., Makridis, C. & Mertens, K. (2020), 'Do Monetary Policy Announcements Shift Household Expectations?', *Federal Reserve Bank of New York Staff Report* (897).

Lorenzoni, G. (2009), 'A Theory of Demand Shocks', *The American Economic Review* **99**(5), 2050–84.

Lorenzoni, G. (2011), 'News and Aggregate Demand Shocks', *The Annual Review of Economics* **3**(1), 537–557.

Loughran, T. & Mcdonald, B. (2016), 'Textual Analysis in Accounting and Finance: A Survey', *Journal of Accounting Research* **54**(4), 1187–1230.

Lucas, R. (1972), 'Expectations and the neutrality of money', *Journal of Economic Theory* **4**(2), 103–124.

Lutkepohl, H. (2007), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag Berlin Heidelberg.

Mankiw, G. N. & Reis, R. (2002), 'Sticky Information versus sticky prices: a proposal to replace the New Keynesian Phillips curve', *Quarterly Journal of Economics* **117**(4), 1295–1328.

McCracken, M. W. & Ng, S. (2015), 'FRED-MD: A Monthly Database for Macroeconomic Research', *Federal Reserve Bank of St. Louis Working Paper* (2015-012B).

Michael D. Bauer and Eric T. Swanson (2020), 'The Fed's Response to Economic News Explains the "Fed Information Effect"', *NBER Working Paper* (27013).

Milani, F. & Treadwell, J. (2012), 'The Effects of Monetary Policy "News" and "Surprises"', *Money, credit and banking* **44**(8), 1667–1692.

Miranda-Agrippino, S. & Ricco, G. (forthcoming), 'The Transmission of Monetary Policy shocks', *American Economic Journal: Macroeconomics* .

Noh, E. (2018), 'Impulse-response analysis with proxy variables', *unpublished* .

Nowzohour, L. & Stracca, L. (2017), 'More than a feeling: confidence, uncertainty and macroeconomic fluctuations', *European Central Bank Working paper series* (No 2100).

Park, T. & Casella, G. (2008), 'The Bayesian Lasso', *Journal of the American Statistical Association* **103**(482), 681–686.

Plagborg-Mollerand, M. & Wolf, C. (2019), 'Local Projections and VARs Estimate the Same Impulse Responses', *Working paper: Department of Economics, Princeton University,* .

Ramey, V. (2016), 'Macroeconomic shocks and their propagation', *Handbook of macroeconomics* **2**, 71–162.

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.

Smith, A. L. & Becker, T. (2015), 'Has Forward Guidance Been Effective?', *Federal Reserve Bank of Kansas City Economic Review* **3**, 57–78.

Soo, C. (2015), 'Quantifying Animal Spirits: News Media and Sentiment in the Housing Market', *Ross School of Business Paper* (1200).

Souleles, N. S. (2004), 'Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys', *Journal of Money, Credit and Banking* **36**(1), 39–72.

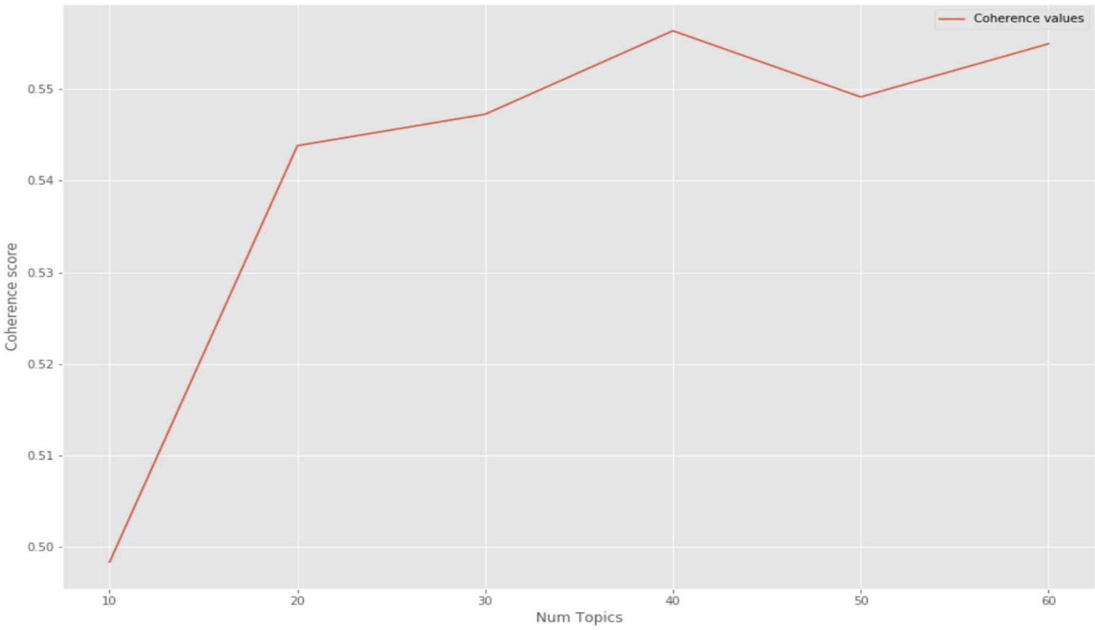
Steinsson, E. N. J. (2019), 'Hight Frequency Identification of Money Non-Neutrality: the Information effect', *NBER Working Paper* (19260).

University of Michigan Survey of Consumers (2019).

URL: *https://data.sca.isr.umich.edu*

Woodford, M. (2001), 'Imperfect common knowledge and the effects of monetary policy', *NBER Working Paper* (8673).

Appendix A. Latent Dirichlet Allocation




```
Num Topics = 10  has Coherence Value of 0.4984
Num Topics = 20  has Coherence Value of 0.5438
Num Topics = 30  has Coherence Value of 0.5473
Num Topics = 40  has Coherence Value of 0.5564
Num Topics = 50  has Coherence Value of 0.5491
Num Topics = 60  has Coherence Value of 0.5549
```

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, commun, 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

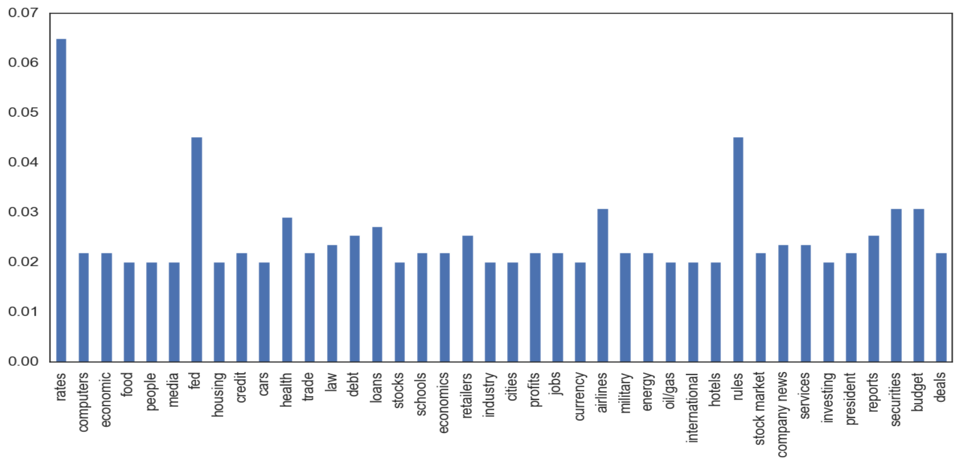


Figure C.1: Topic proportions for the paragraph 1

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

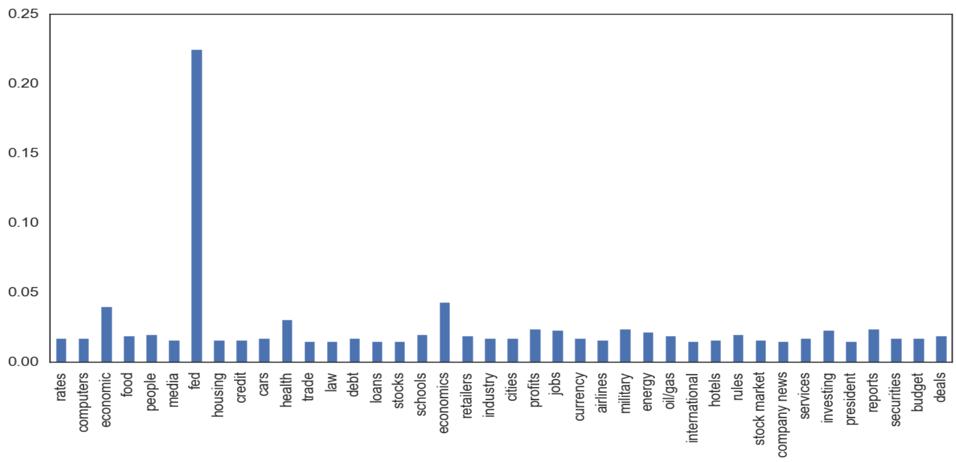


Figure C.2: Topic proportions for the paragraph 2

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

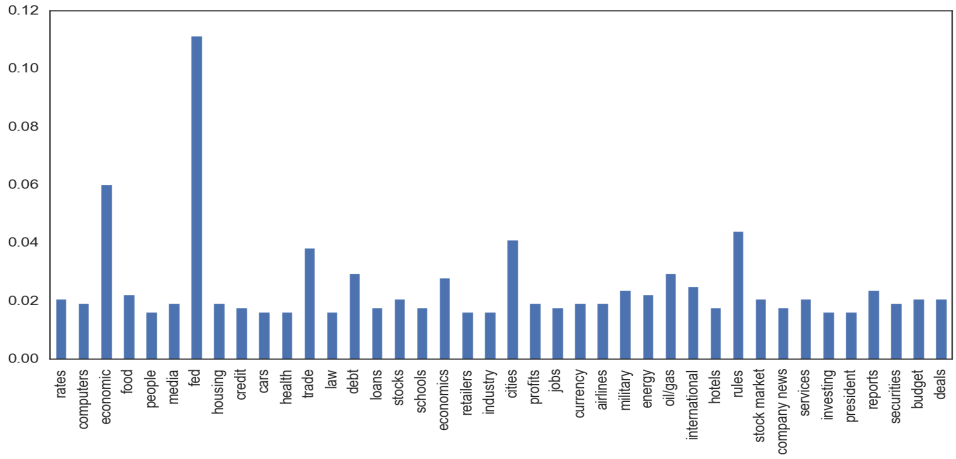


Figure C.3: Topic proportions for the paragraph 3

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

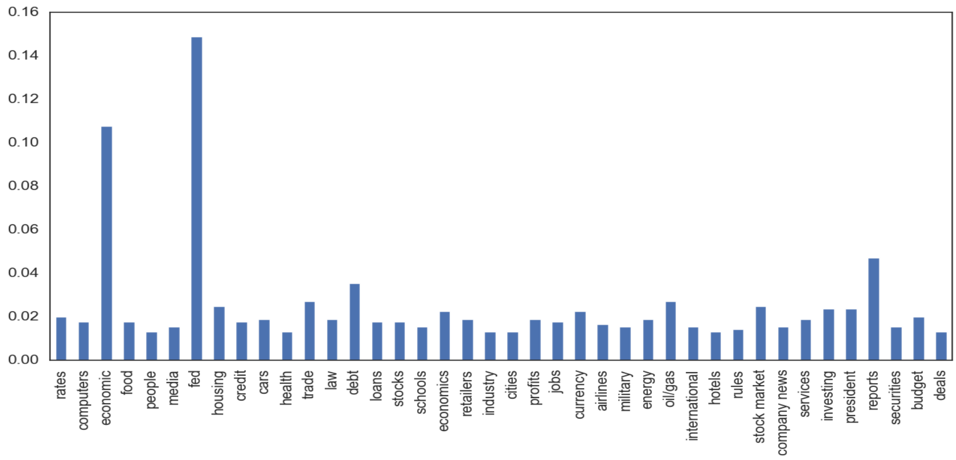


Figure C.4: Topic proportions for the paragraph 4

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

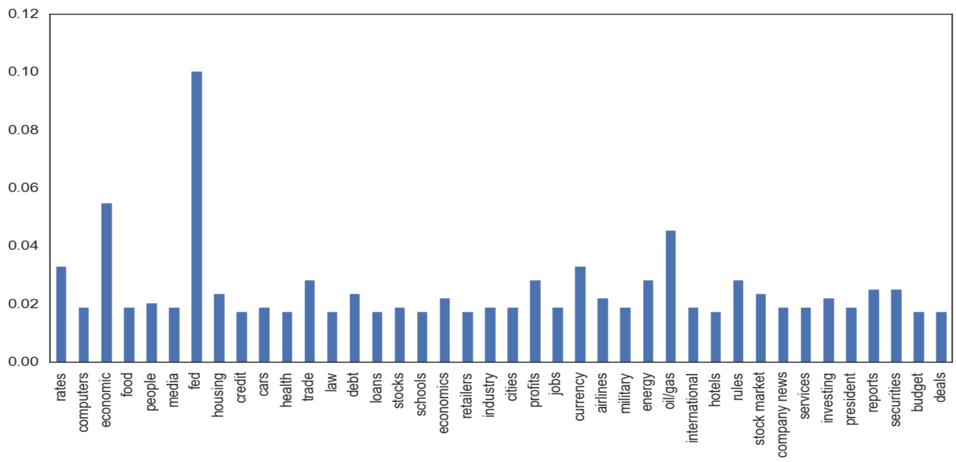


Figure C.5: Topic proportions for the paragraph 5

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

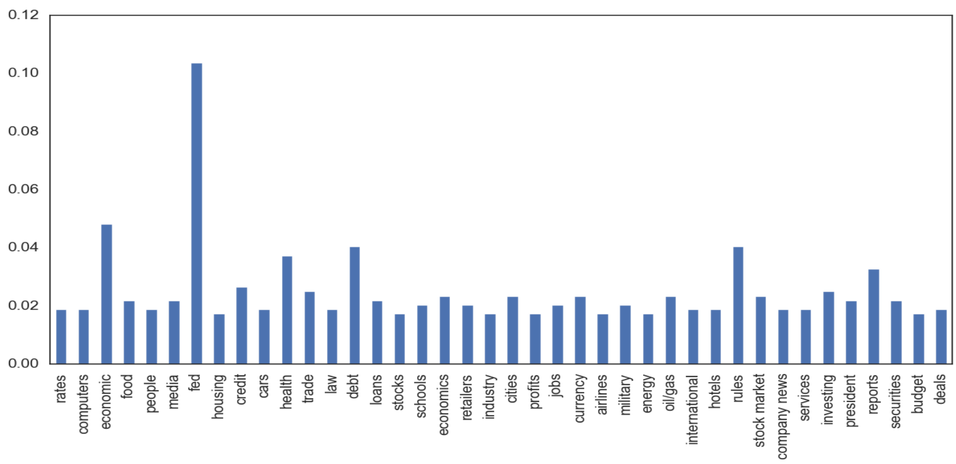


Figure C.6: Topic proportions for the paragraph 6

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

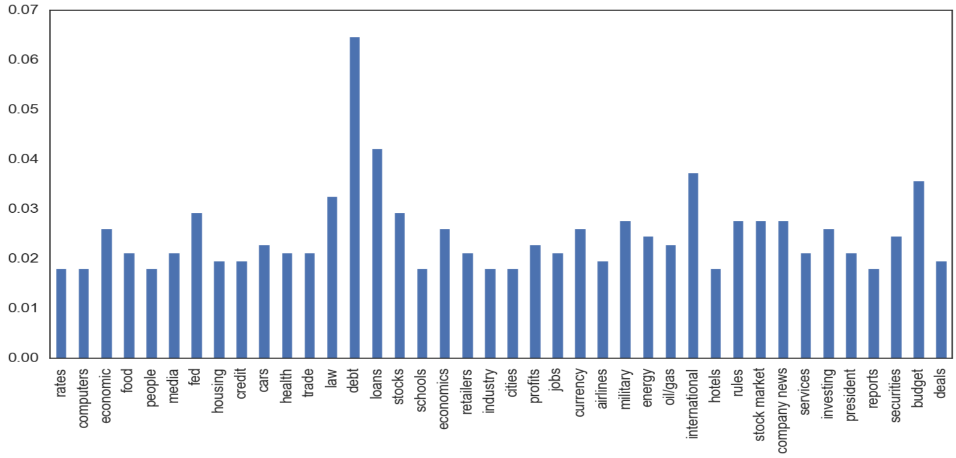


Figure C.7: Topic proportions for the paragraph 7

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

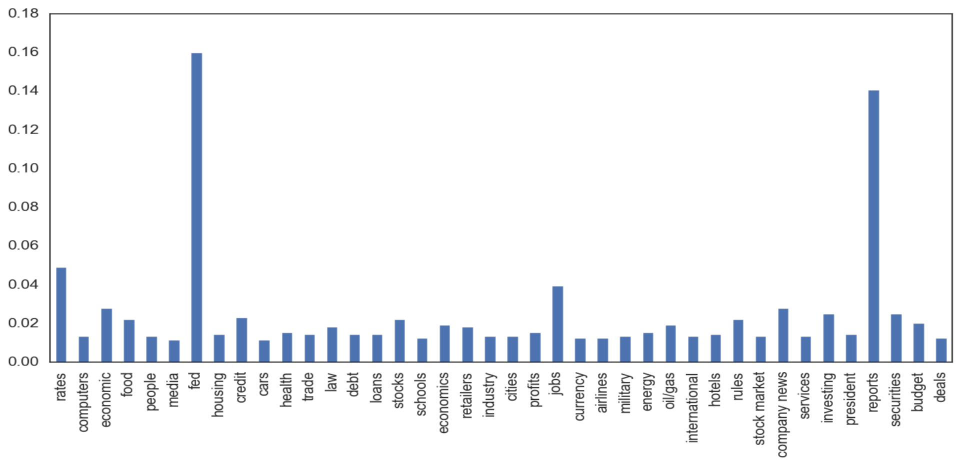


Figure C.8: Topic proportions for the paragraph 8

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

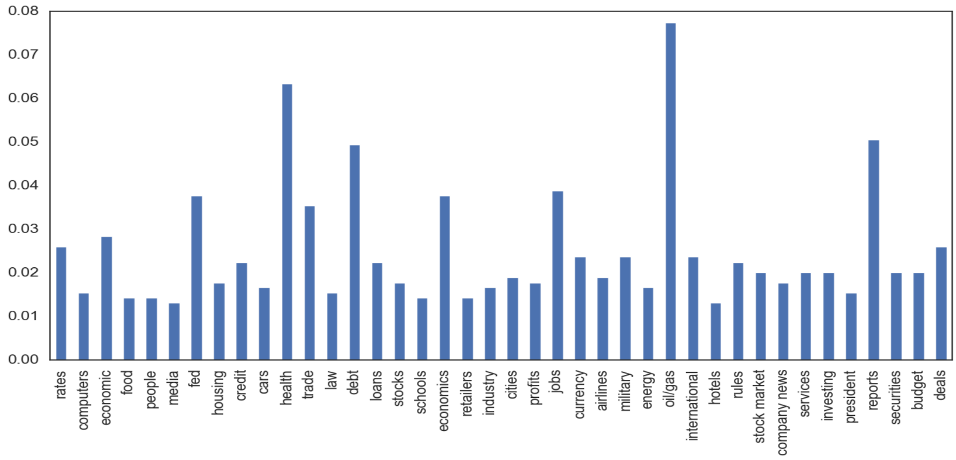


Figure C.9: Topic proportions for the paragraph 9

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

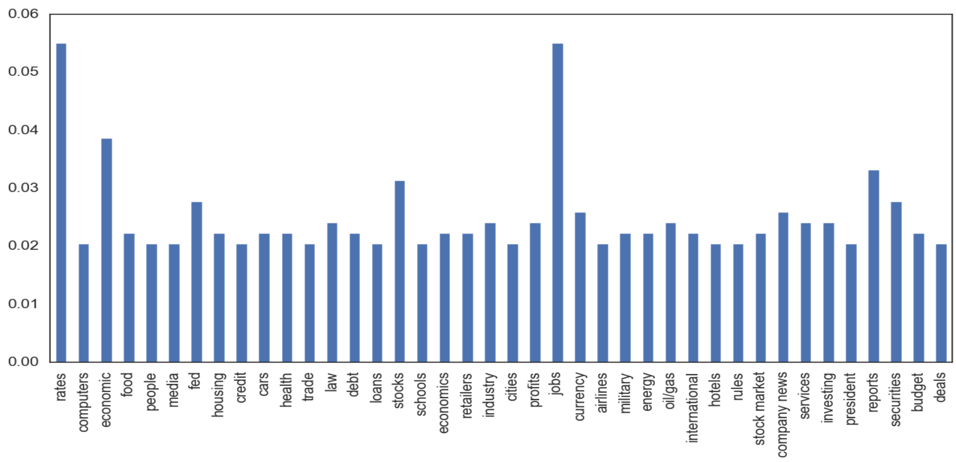


Figure C.10: Topic proportions for the sentence 1

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

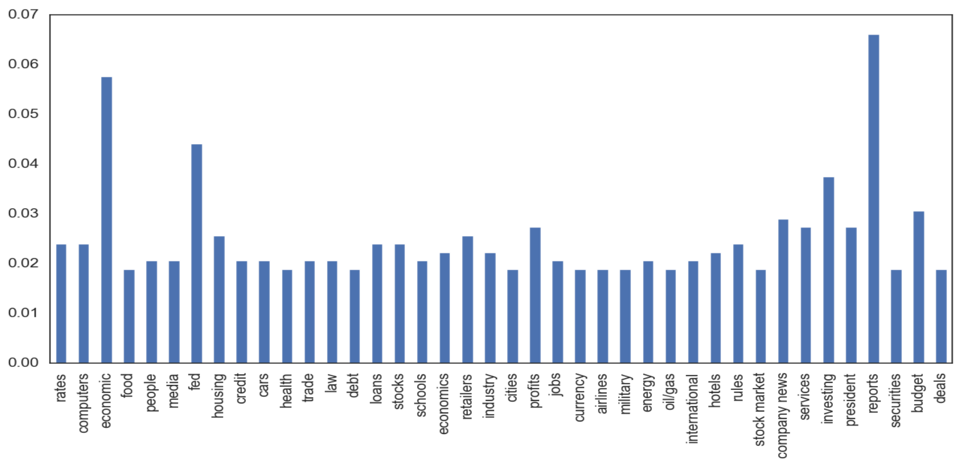


Figure C.11: Topic proportions for the sentence 2

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

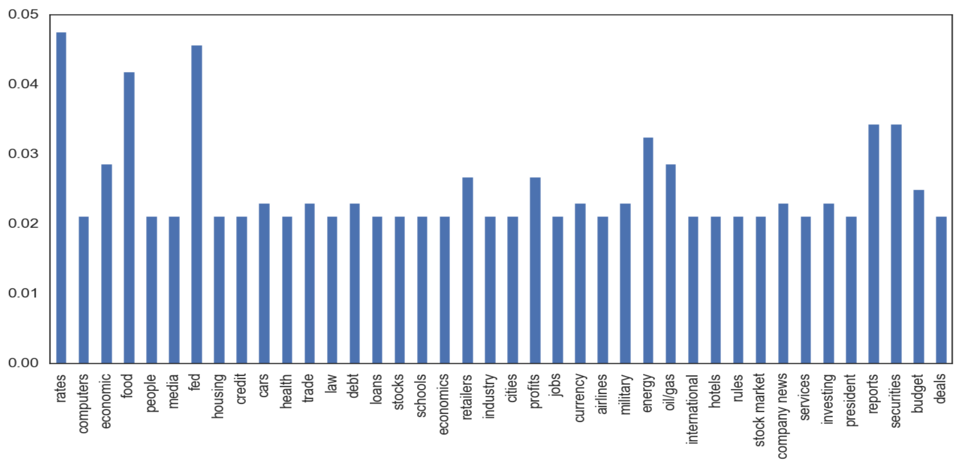


Figure C.12: Topic proportions for the sentence 3

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

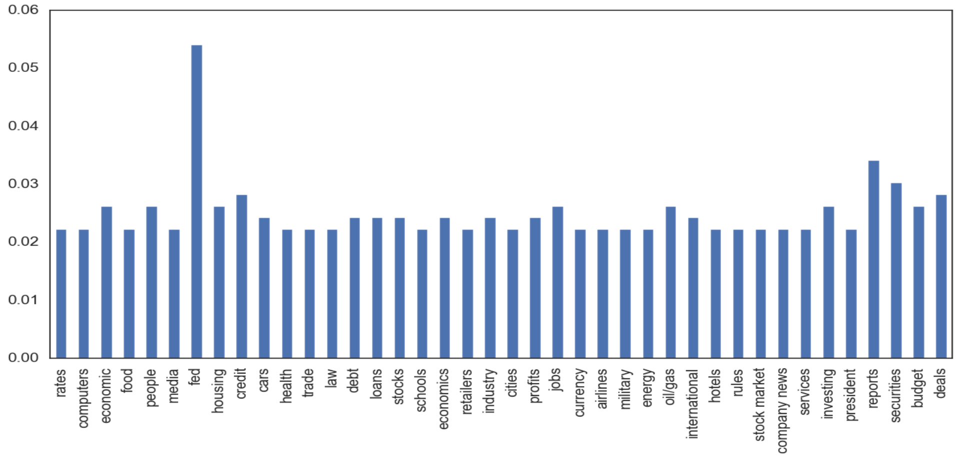


Figure C.13: Topic proportions for the sentence 4

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

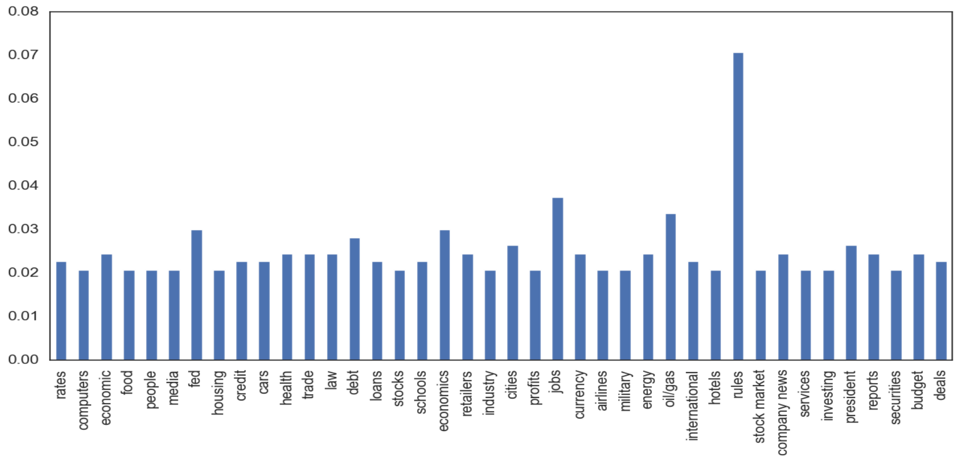


Figure C.14: Topic proportions for the sentence 5

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

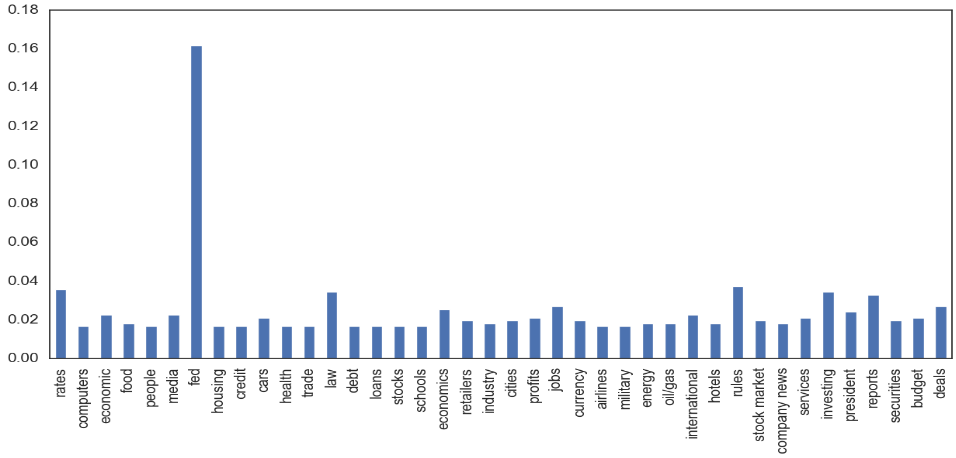


Figure C.15: Topic proportions for the sentence 6

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

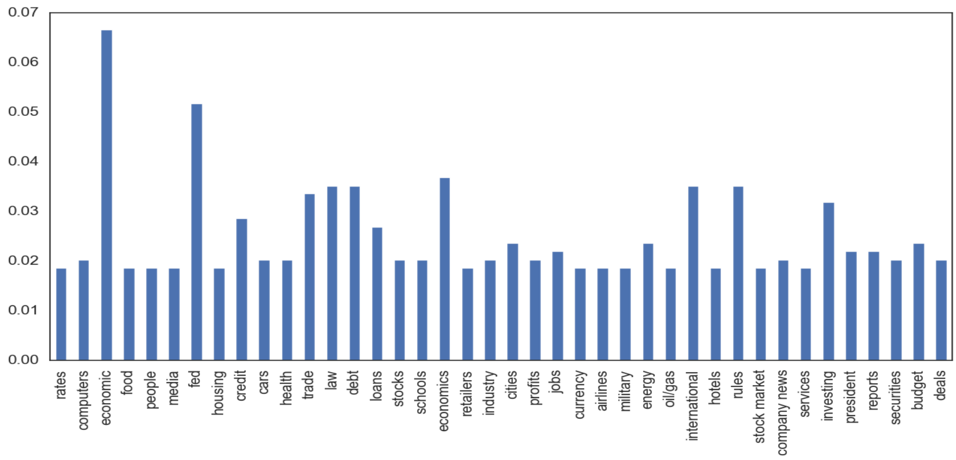


Figure C.16: Topic proportions for the sentence 7

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

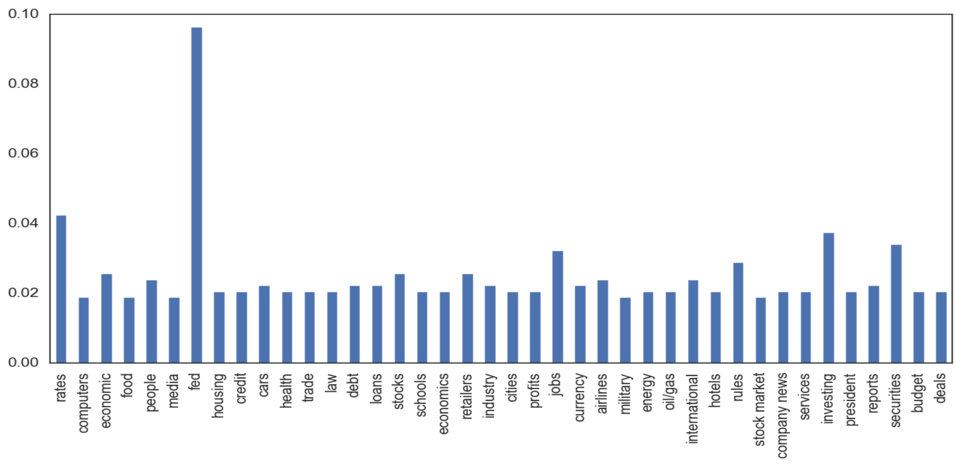


Figure C.17: Topic proportions for the sentence 8

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

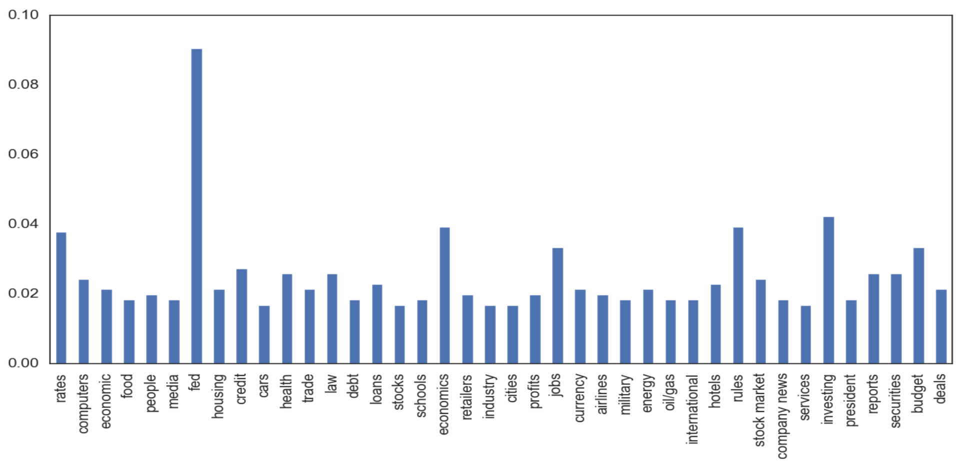


Figure C.18: Topic proportions for the sentence 9

C.3 Classifying QE sentences

1. As previously announced, over the next few quarters the federal reserve will purchase large quantities of agency debt and mortgage-backed securities to provide support to the mortgage and housing markets, and it stands ready to expand its purchases of agency debt and mortgage-backed securities as conditions warrant. 2008-12-16

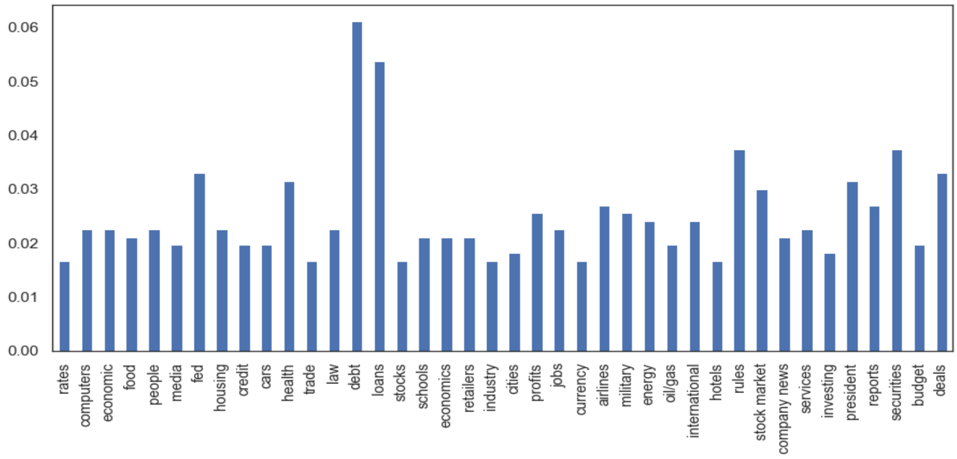


Figure C.19: Topic proportions for the sentence 1

2. The committee also is prepared to purchase longer-term treasury securities if evolving circumstances indicate that such transactions would be particularly effective in improving conditions in private credit markets. 2009-01-28

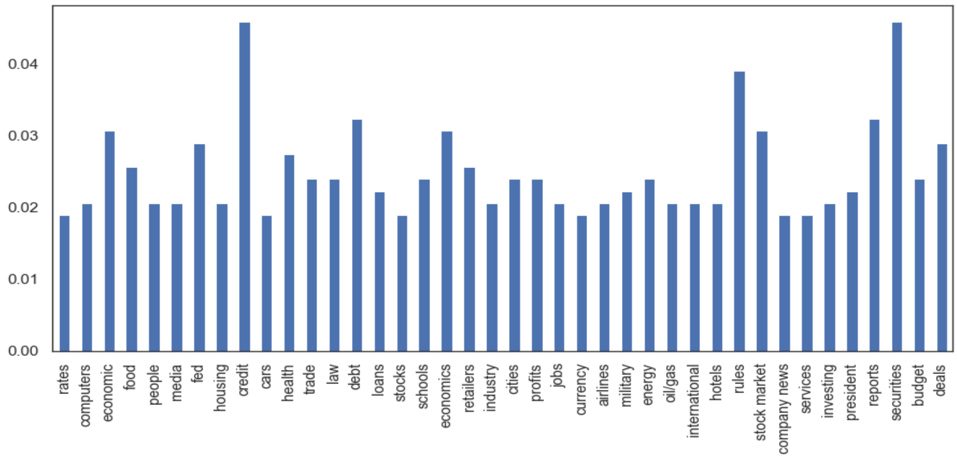


Figure C.20: Topic proportions for the sentence 2

3. To provide greater support to mortgage lending and housing markets, the committee decided today to increase the size of the federal reserve?s balance sheet further by purchasing up to an additional $750 billion of agency mortgage-backed securities, bringing its total purchases of these securities to up to. 2009-03-18

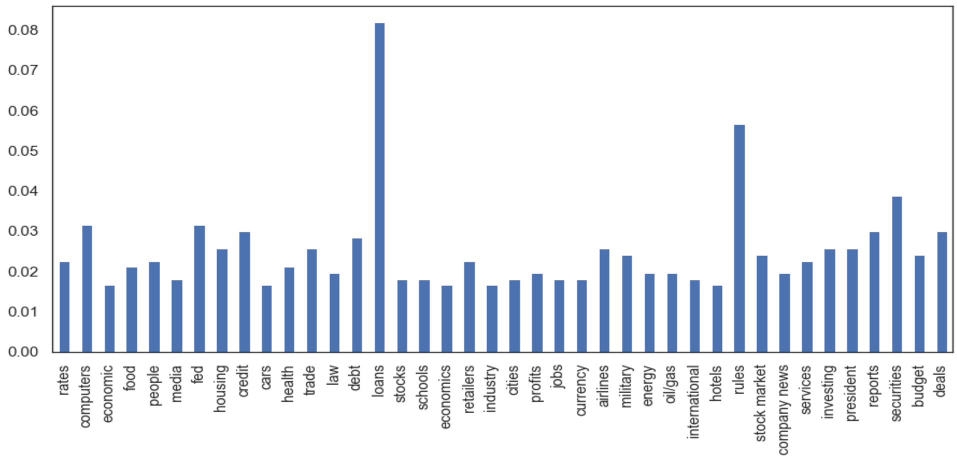


Figure C.21: Topic proportions for the sentence 3

4. 125 trillion this year, and to increase its purchases of agency debt this year by up to \$100 billion to a total of up to \$200 billion

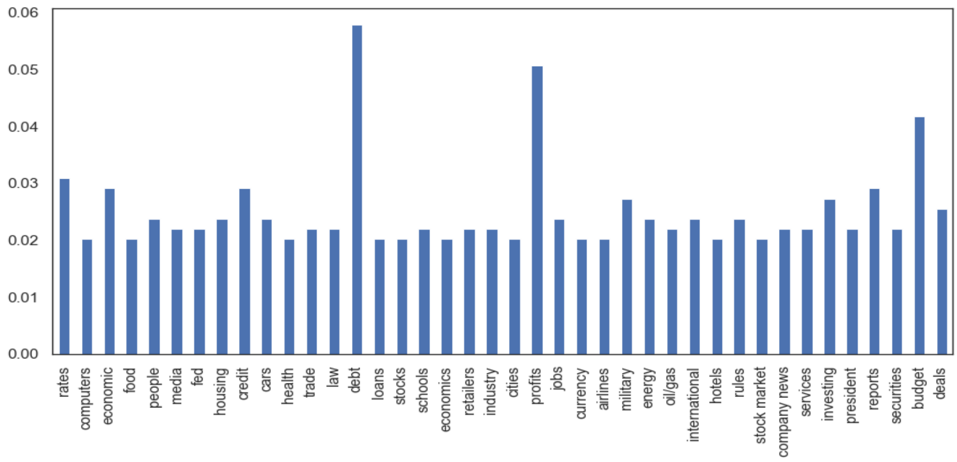


Figure C.22: Topic proportions for the sentence 4

5. in addition, the committee intends to purchase a further \$600 billion of longer-term treasury securities by the end of the second quarter of 2011, a pace of about \$75 billion per month. 2010-11-03

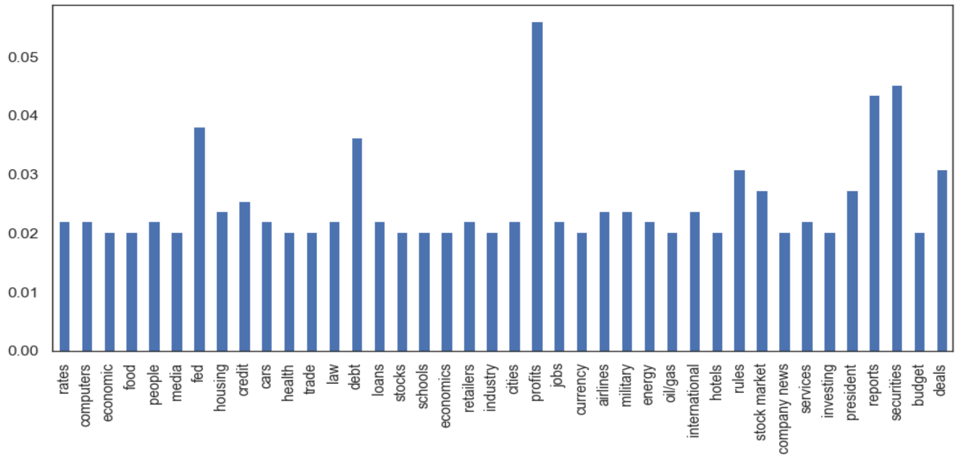


Figure C.23: Topic proportions for the sentence 5

6. The committee will closely monitor incoming information on economic and financial developments in coming months and will continue its purchases of treasury and agency mortgage-backed securities, and employ its other policy tools as appropriate, until the outlook for the labor market has improved substantially in a context of price stability. 2014-09-17

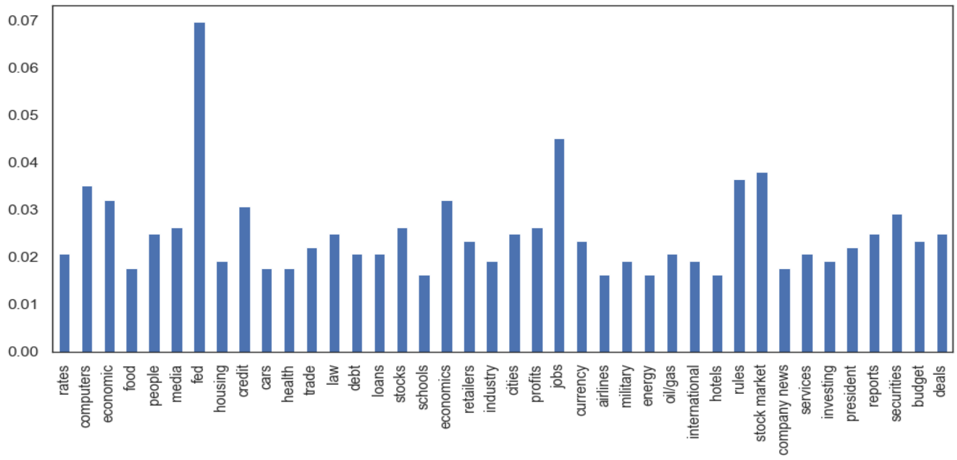


Figure C.24: Topic proportions for the sentence 6

7. In addition, the federal reserve will conduct term and overnight repurchase agreement operations at least through january of next year to ensure that the supply of reserves remains ample even during periods of sharp increases in non-reserve liabilities, and to mitigate the risk of money market pressures that could adversely affect policy implementation. 2019-10-11

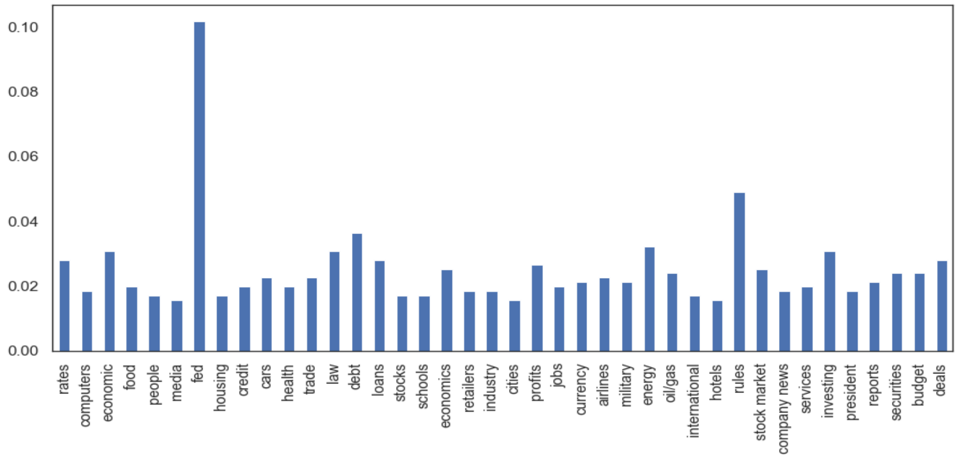


Figure C.25: Topic proportions for the sentence 7

C.4 Classifying Forward Guidance sentences

1. the committee anticipates, based on its current assessment, 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 this month, 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. 2014-10-29

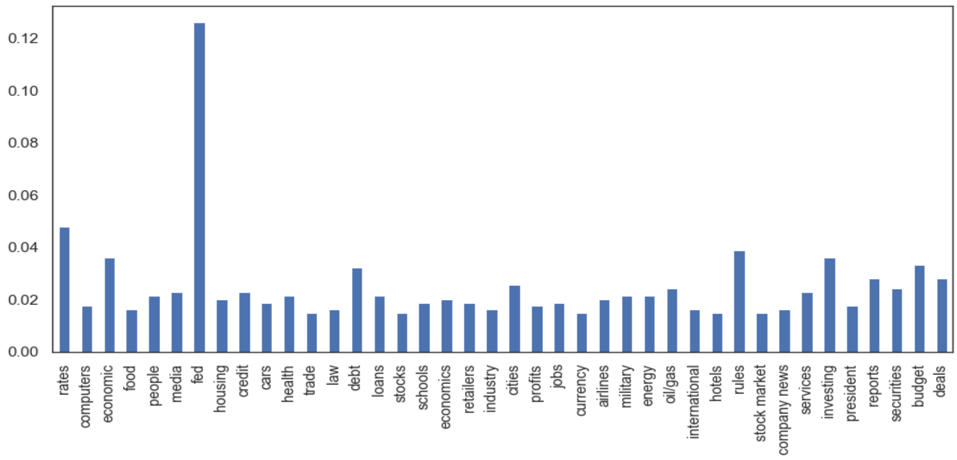


Figure C.26: Topic proportions for the sentence 1

2. 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. 2014-12-17

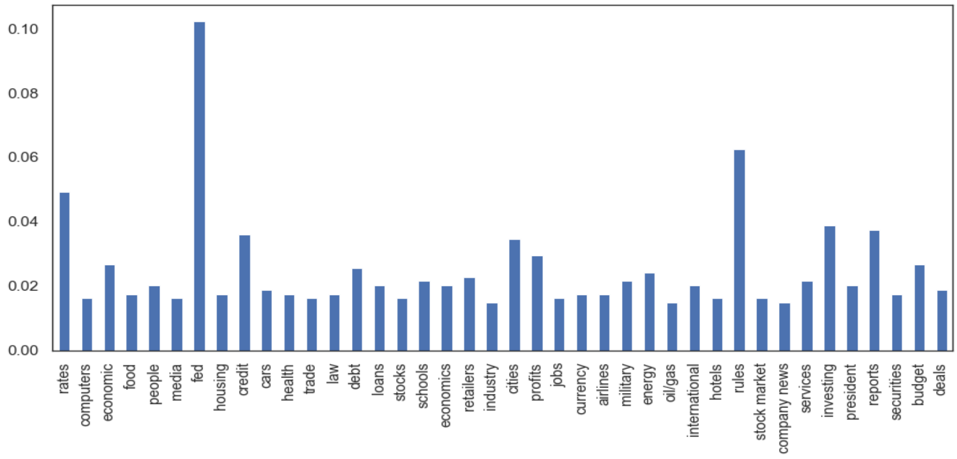


Figure C.27: Topic proportions for the sentence 2

3. The committee continues to anticipate, based on its assessment of these factors, that it likely will be appropriate to maintain the current target range for the federal funds rate well past the time that the unemployment rate declines below 6-1/2 percent, especially if projected inflation continues to run below the committee's 2 percent longer-run goal. 2014-01-29

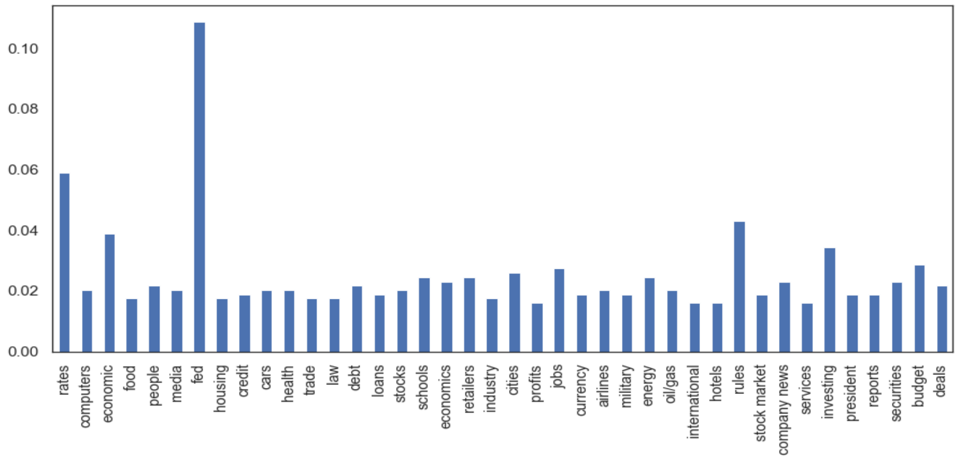


Figure C.28: Topic proportions for the sentence 3

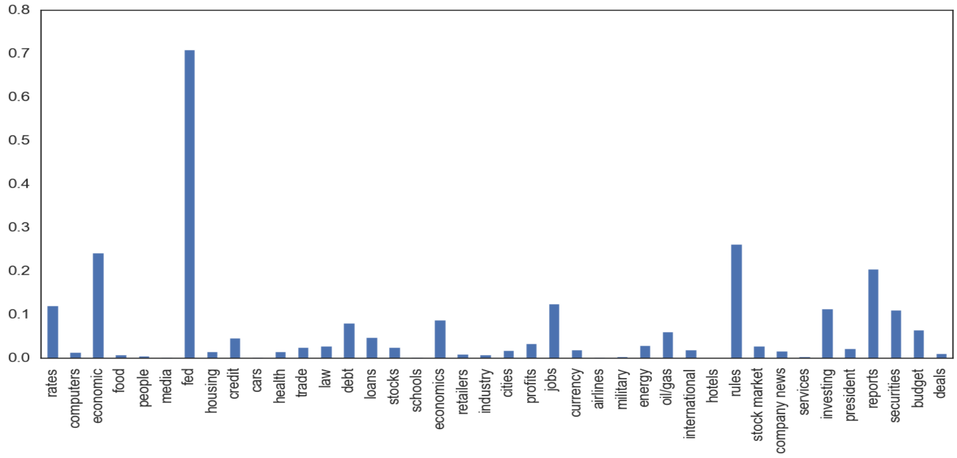


Figure C.29: Aggregated topic proportions by sentence

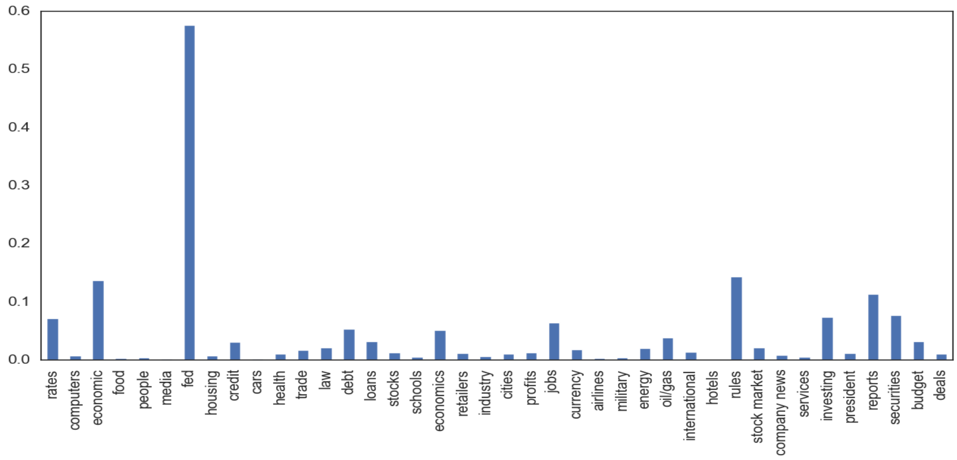


Figure C.30: Aggregated topic proportions by paragraph

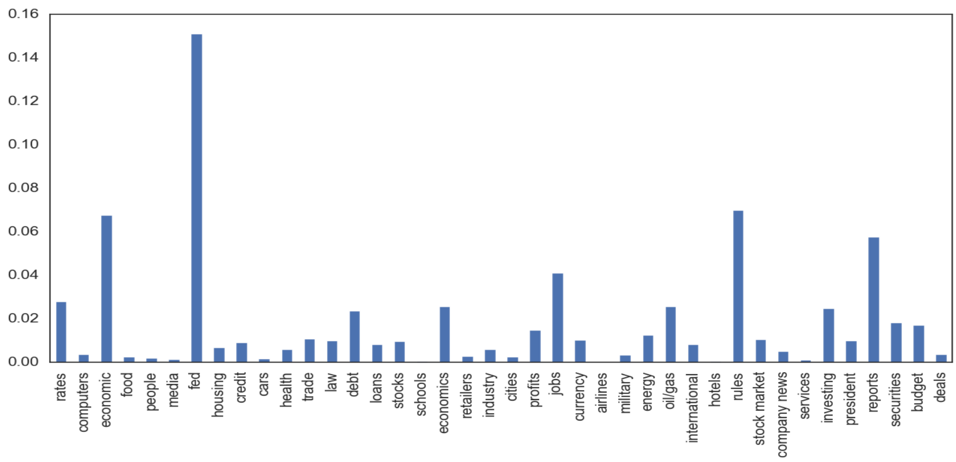


Figure C.31: Aggregated topic proportions by sentence with sign adjustment

Appendix D. Information in FOMC statements

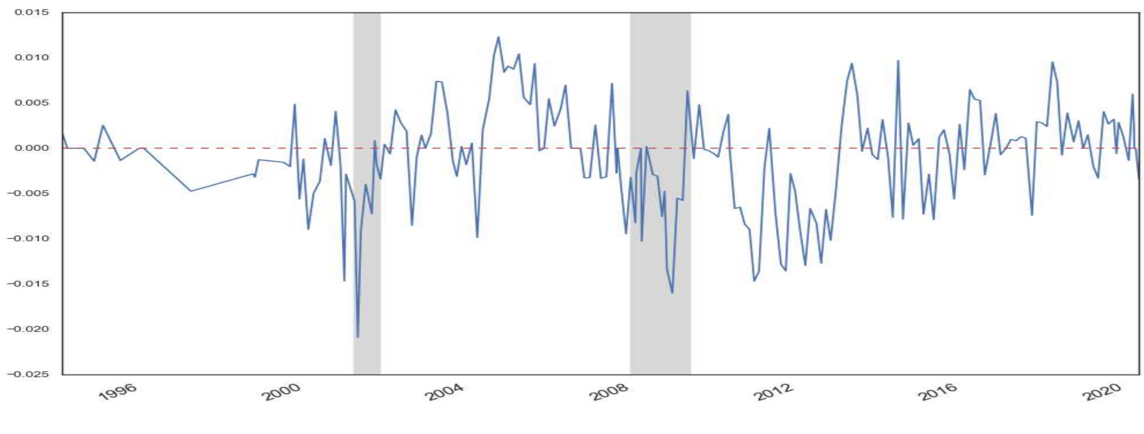


Figure D.1: Economic topic

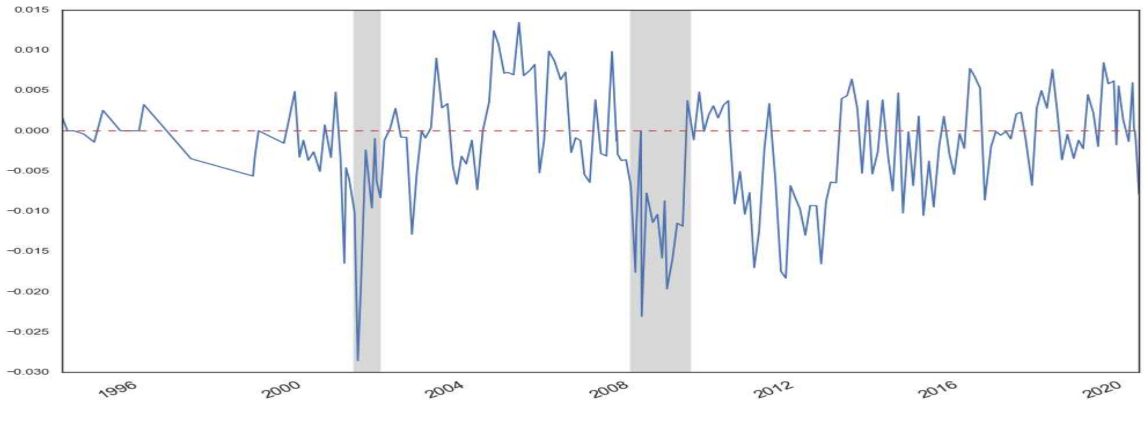


Figure D.2: Economic topic from combination of dictionaries without directional words

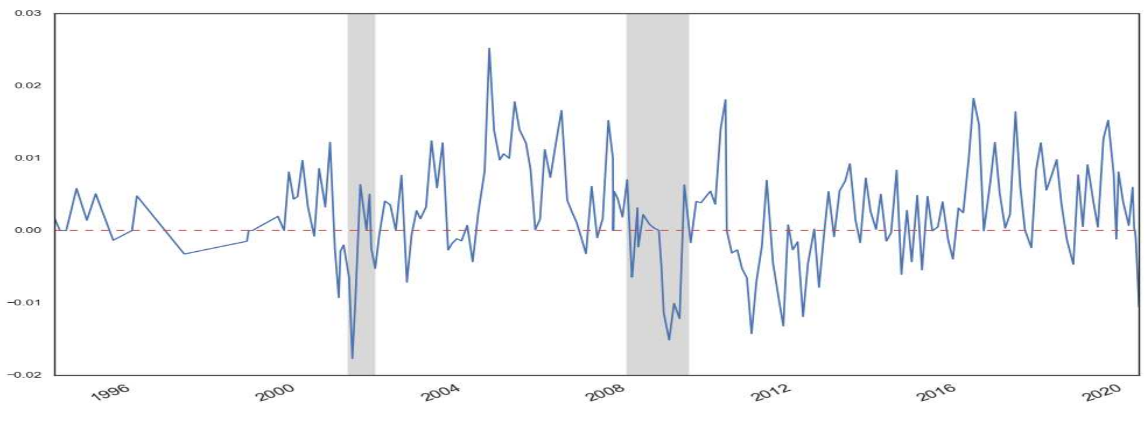


Figure D.3: Economic topic from combination of dictionaries with directional words

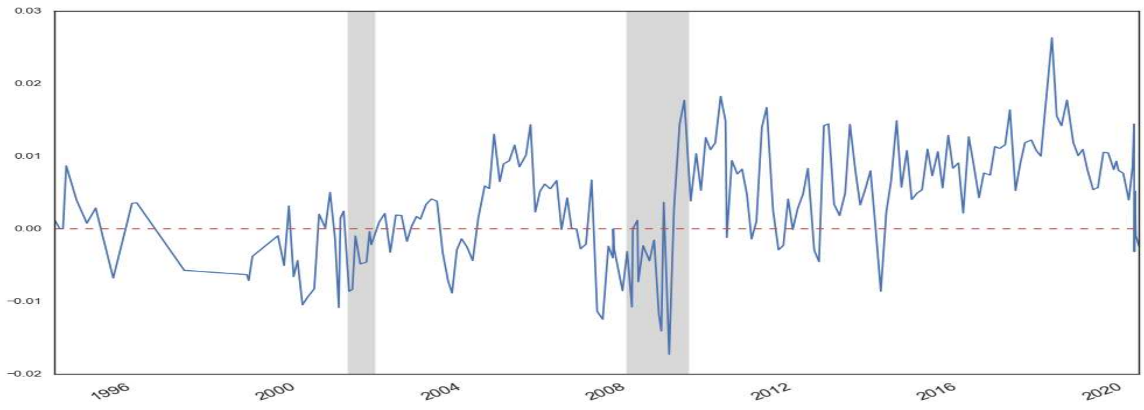


Figure D.4: Fed topic

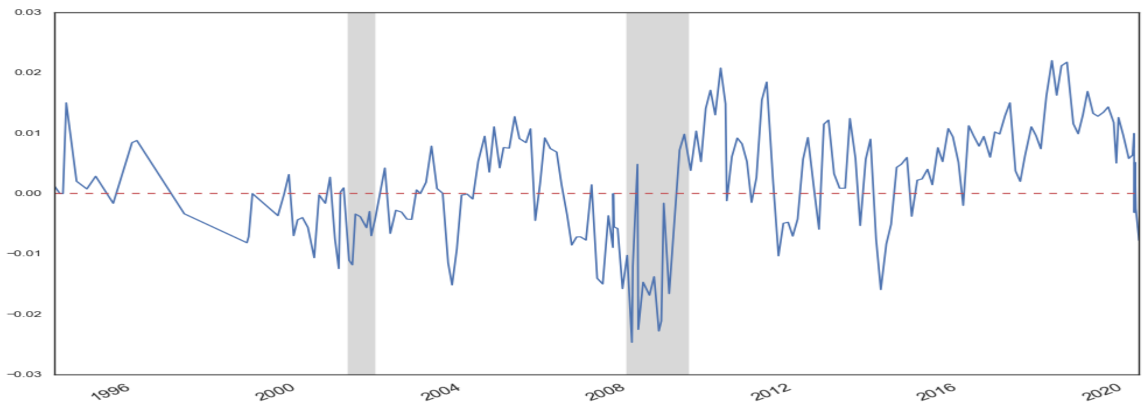


Figure D.5: Fed topic from combination of dictionaries without directional words

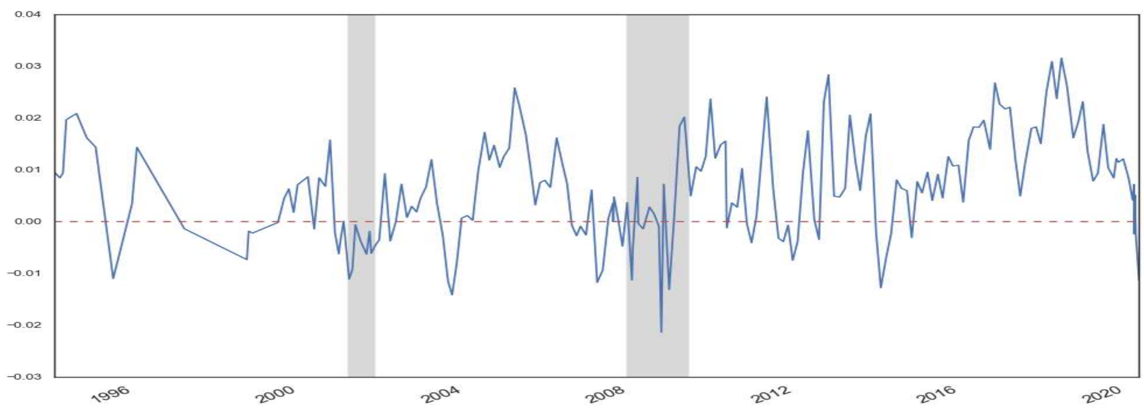


Figure D.6: Fed topic from combination of dictionaries with directional words

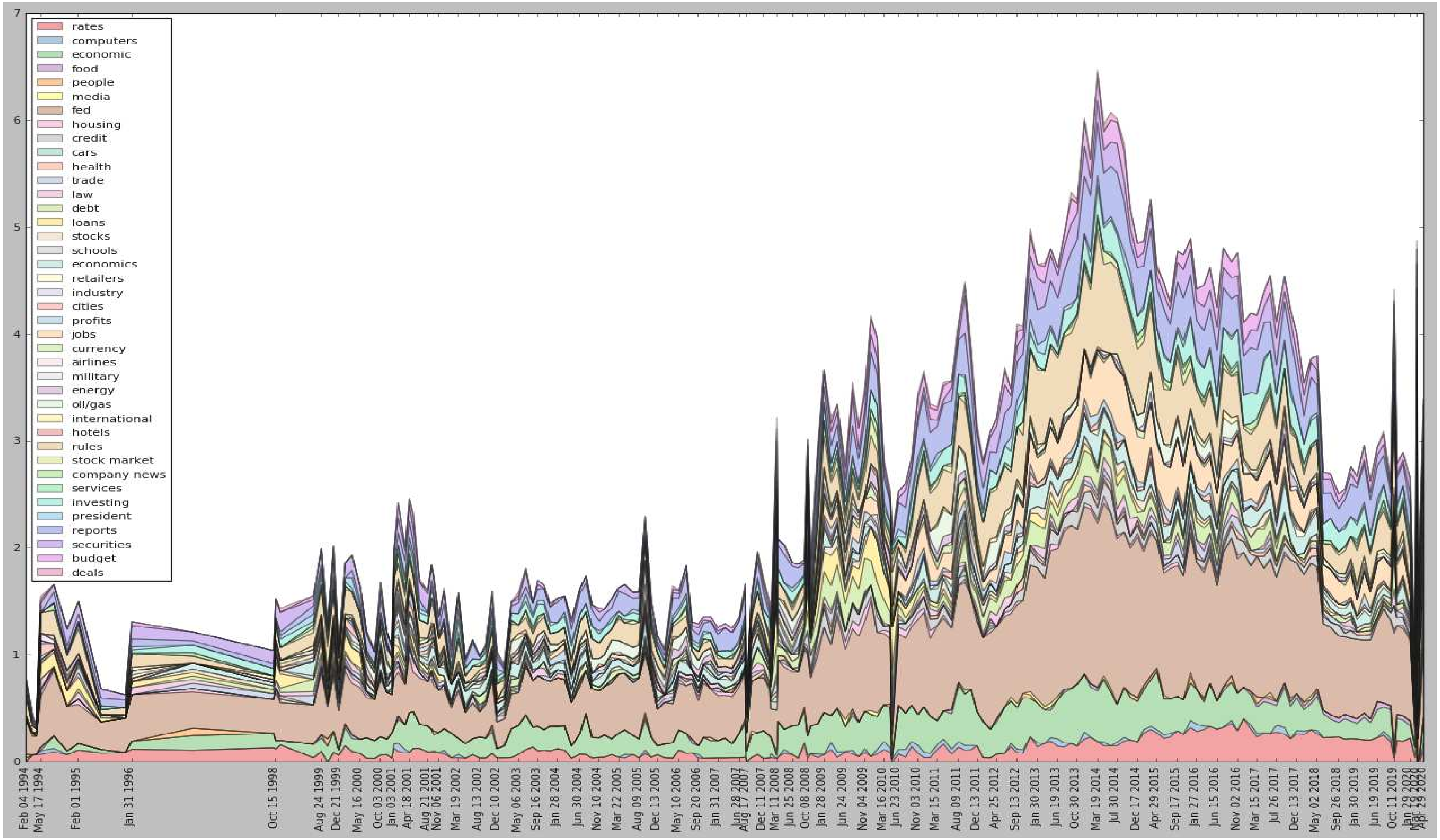


Figure D.7: Topic frequencies over time

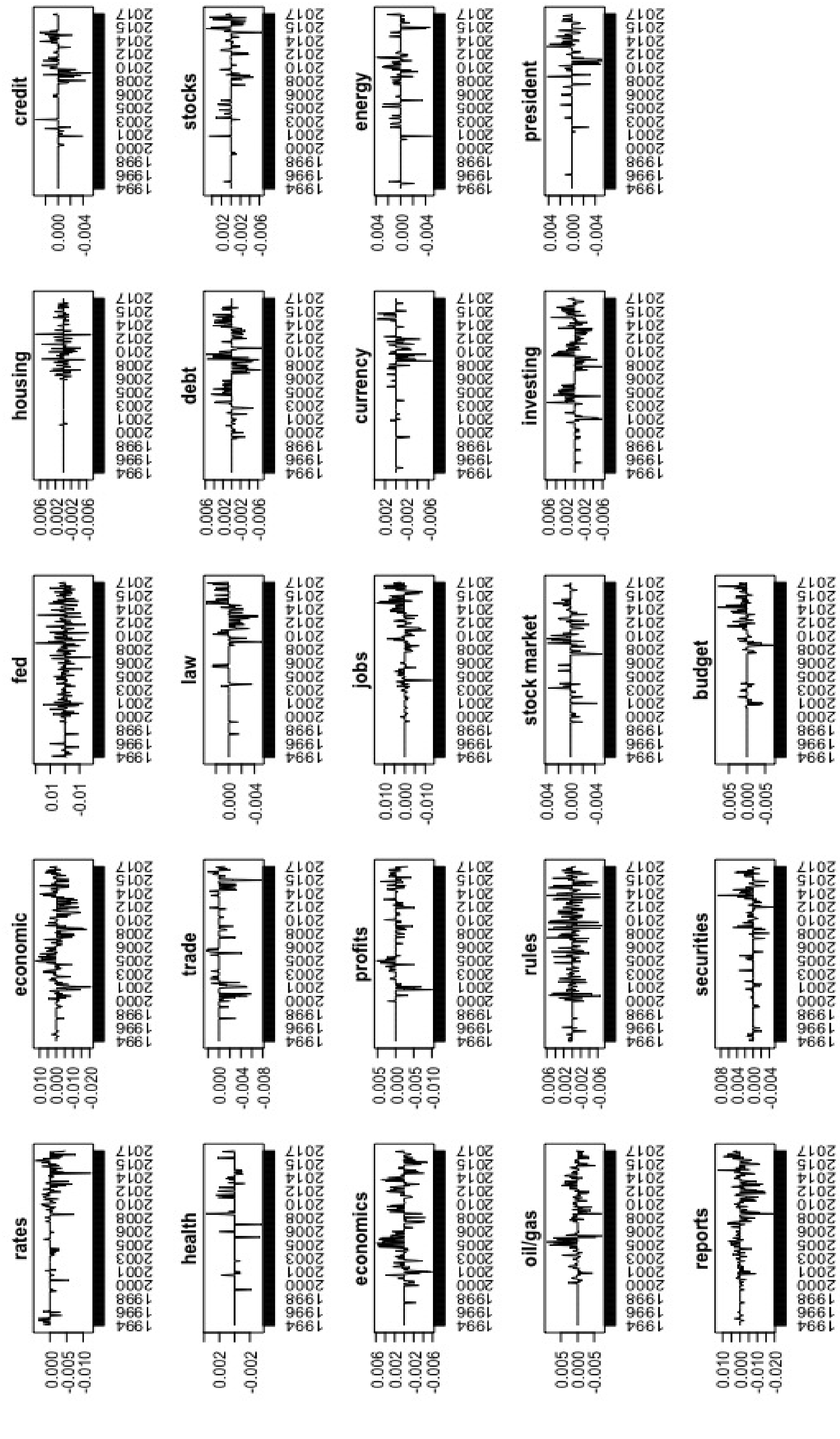


Figure D.8: Topic time series positiveness tone adjusted

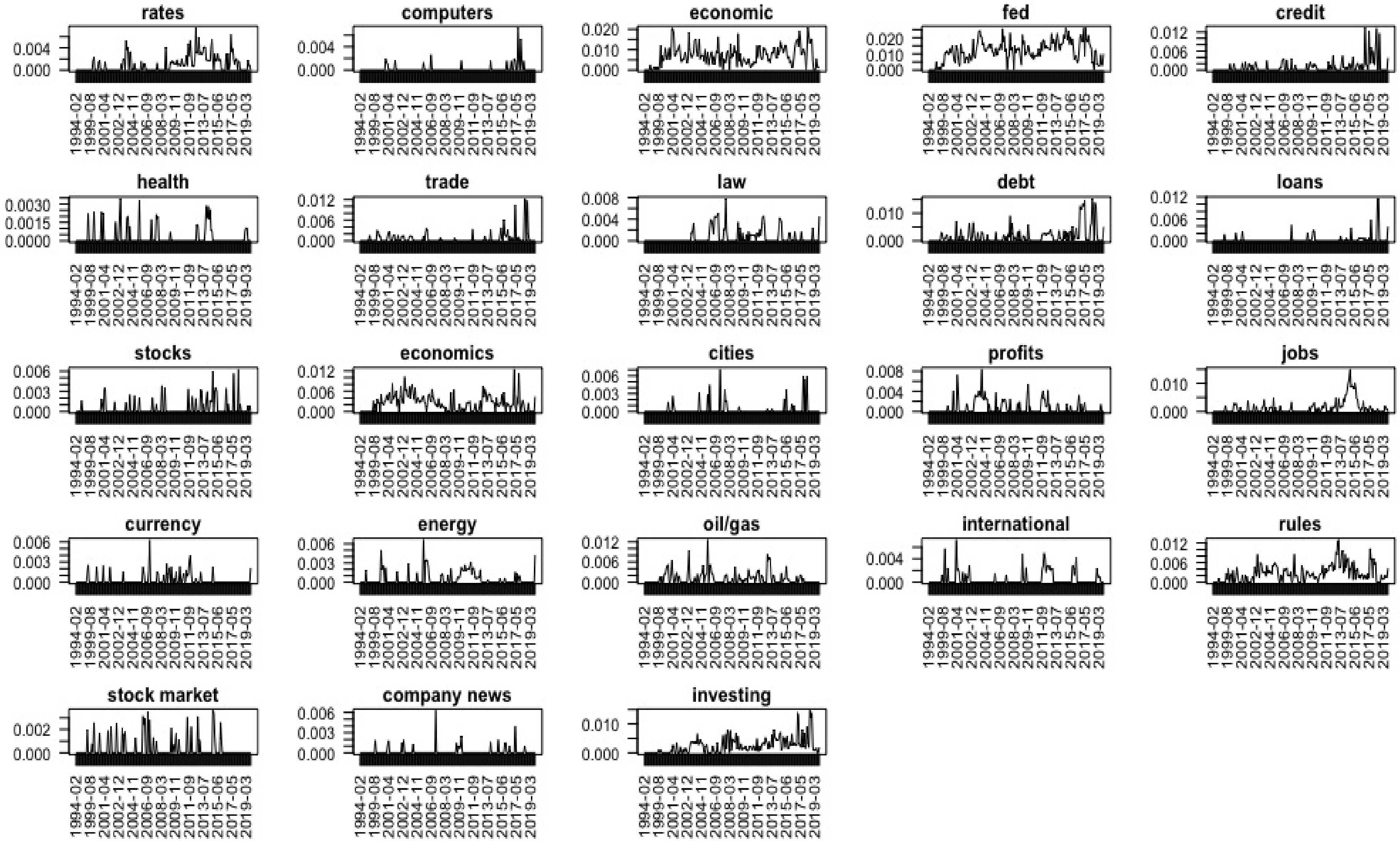
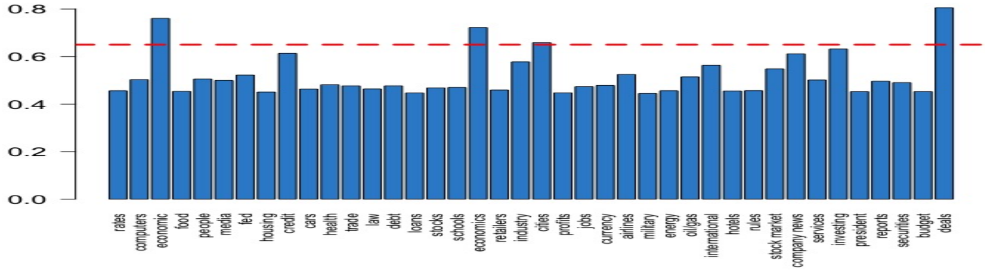
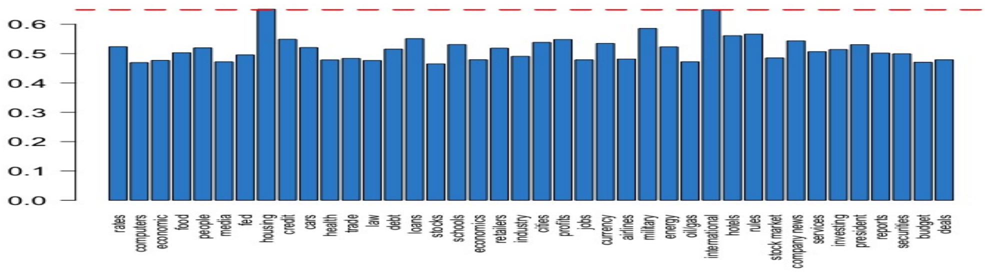


Figure D.9: Topic time series uncertainty tone adjusted

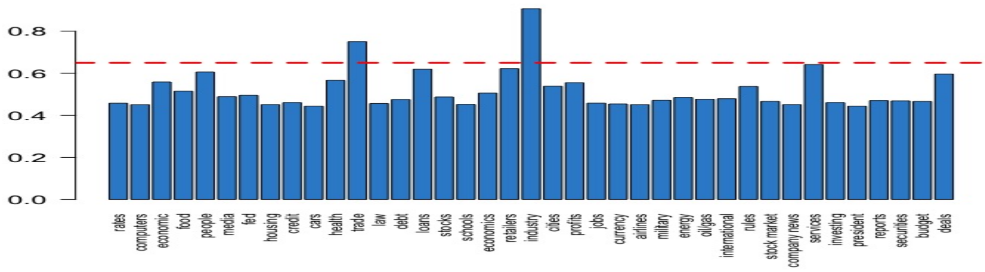
Appendix E. Model selection



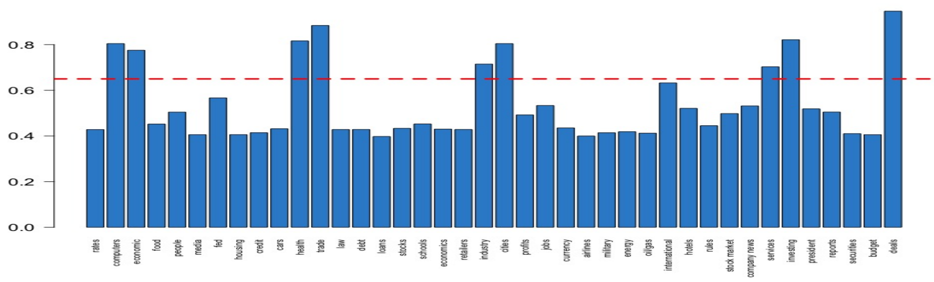
(a) Sign adjustment



(b) Non-adjusted frequency



(c) Uncertainty



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

Figure E.1: Posterior inclusion probabilities

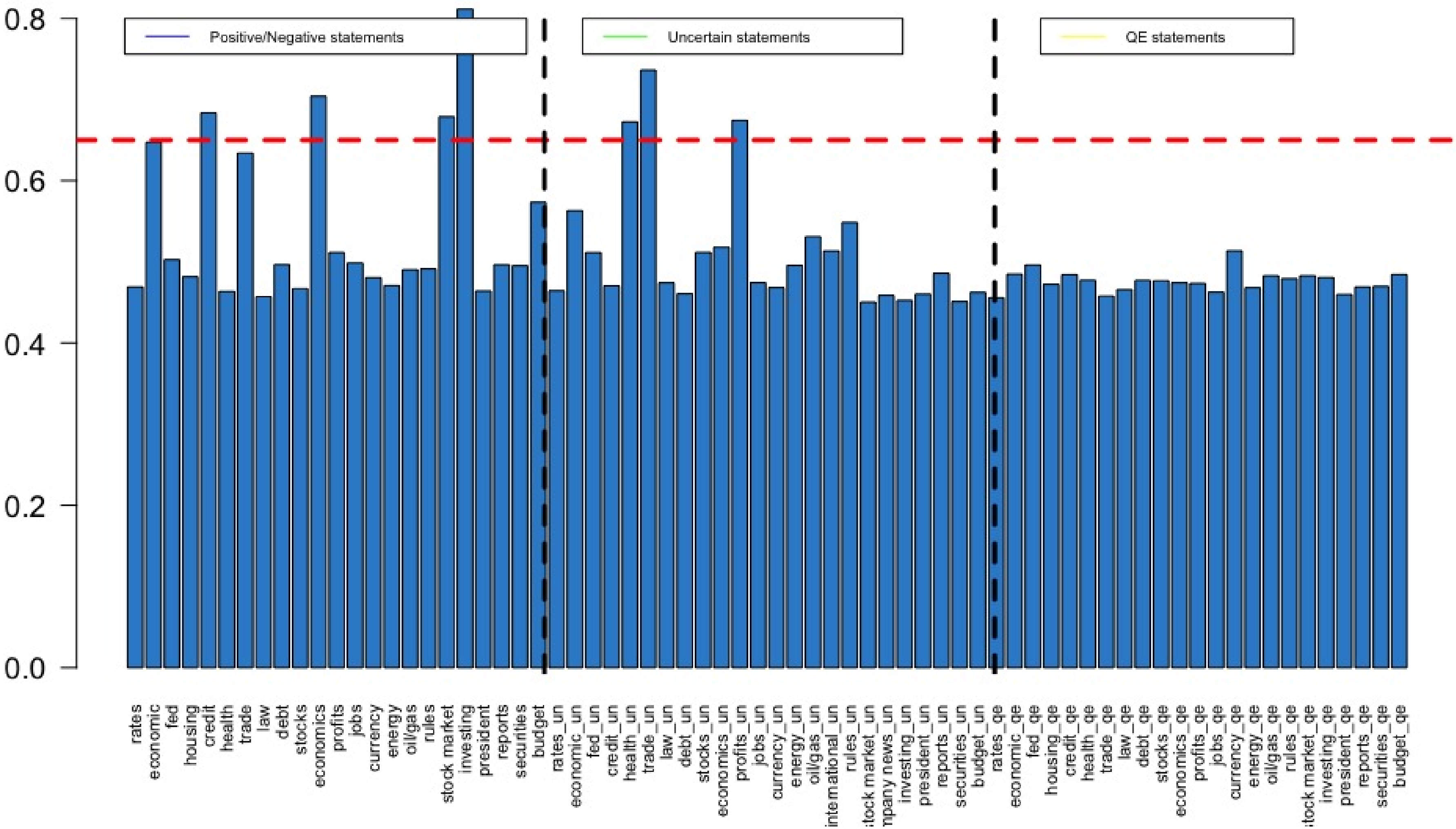


Figure E.2: Bayesian Lasso for surprises in 3m federal funds futures

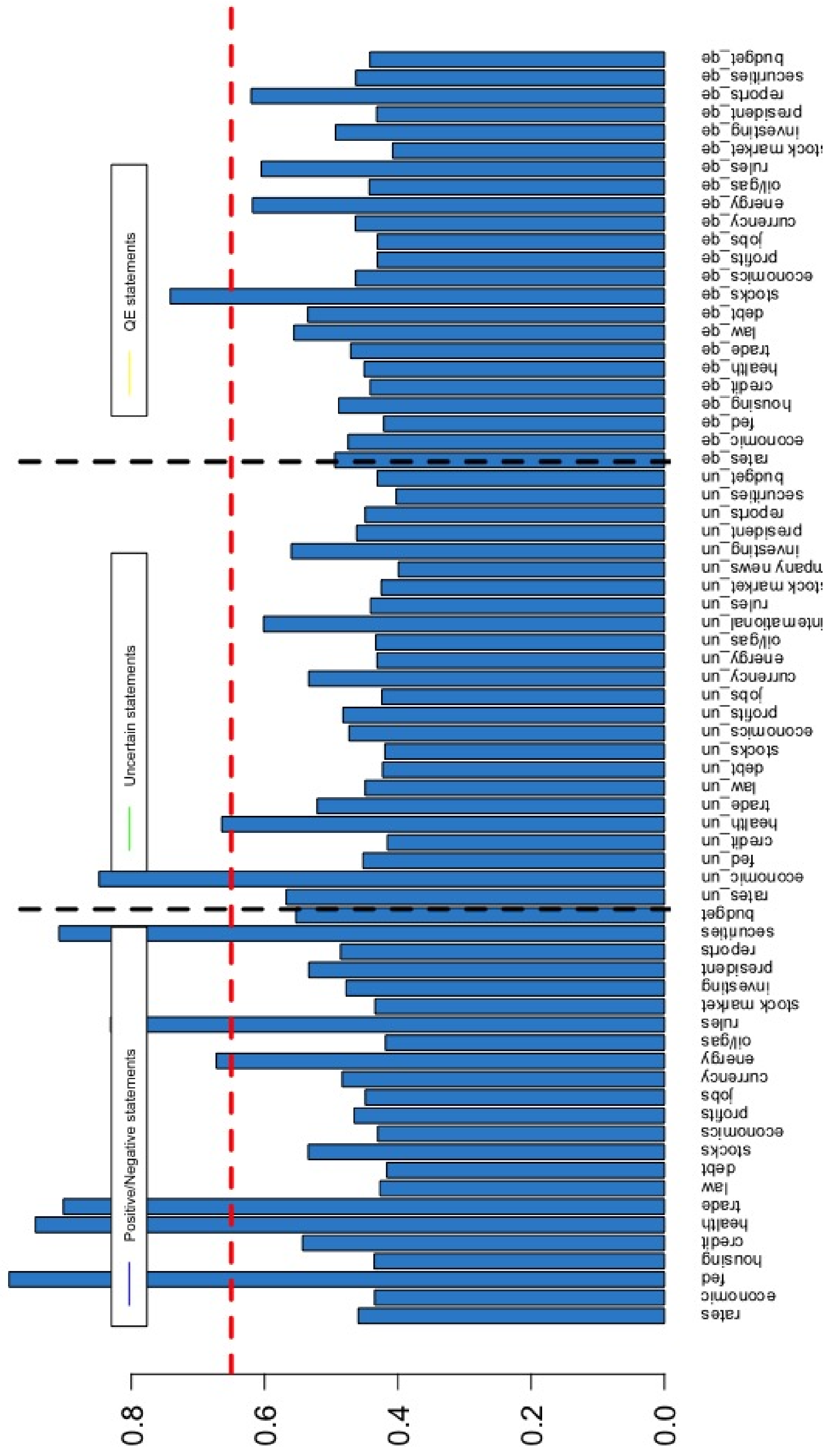


Figure E.3: Bayesian Lasso for surprises in S&P

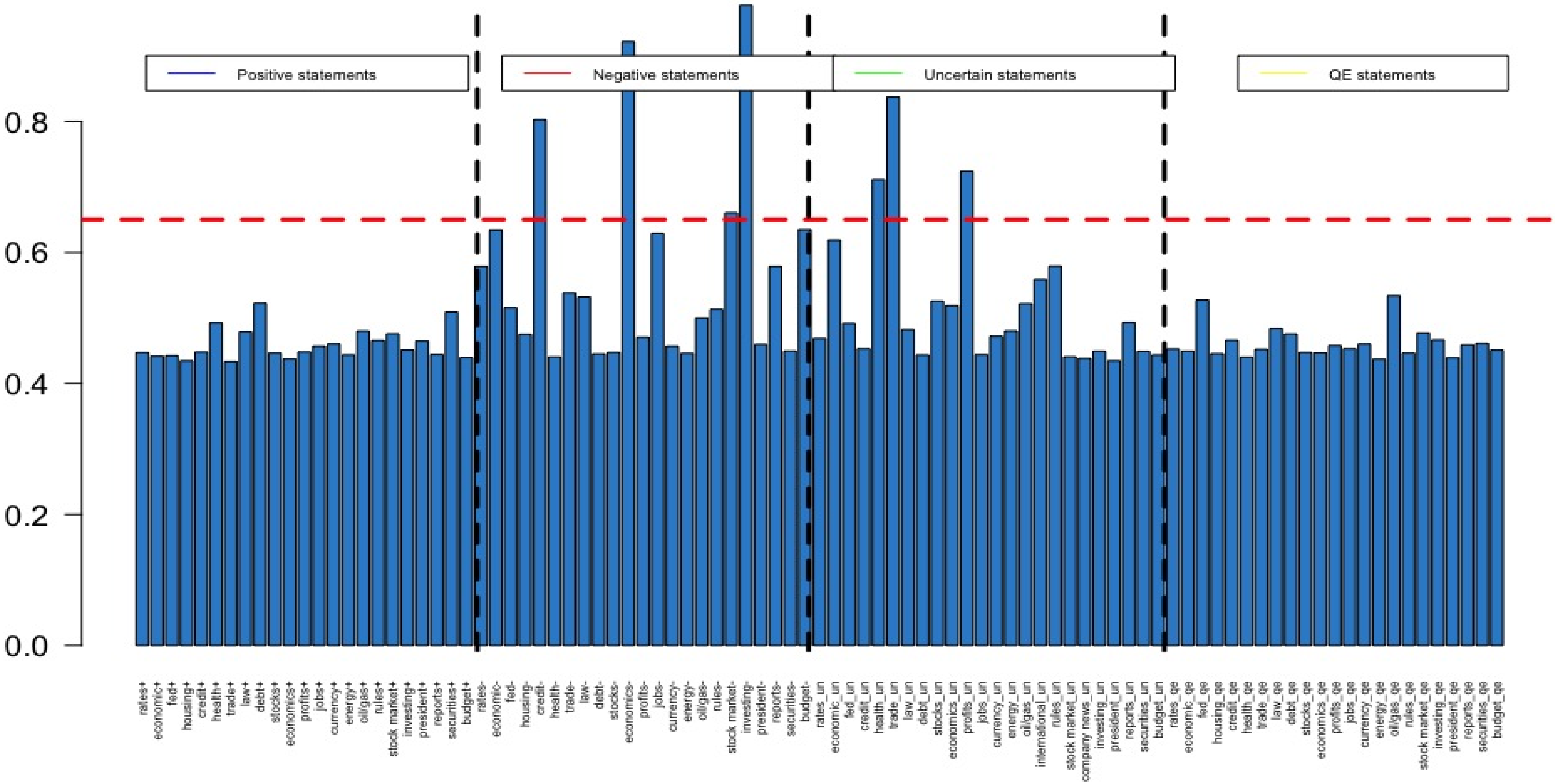


Figure E.4: Bayesian Lasso for surprises in n 3m federal funds futures with asymmetric effect

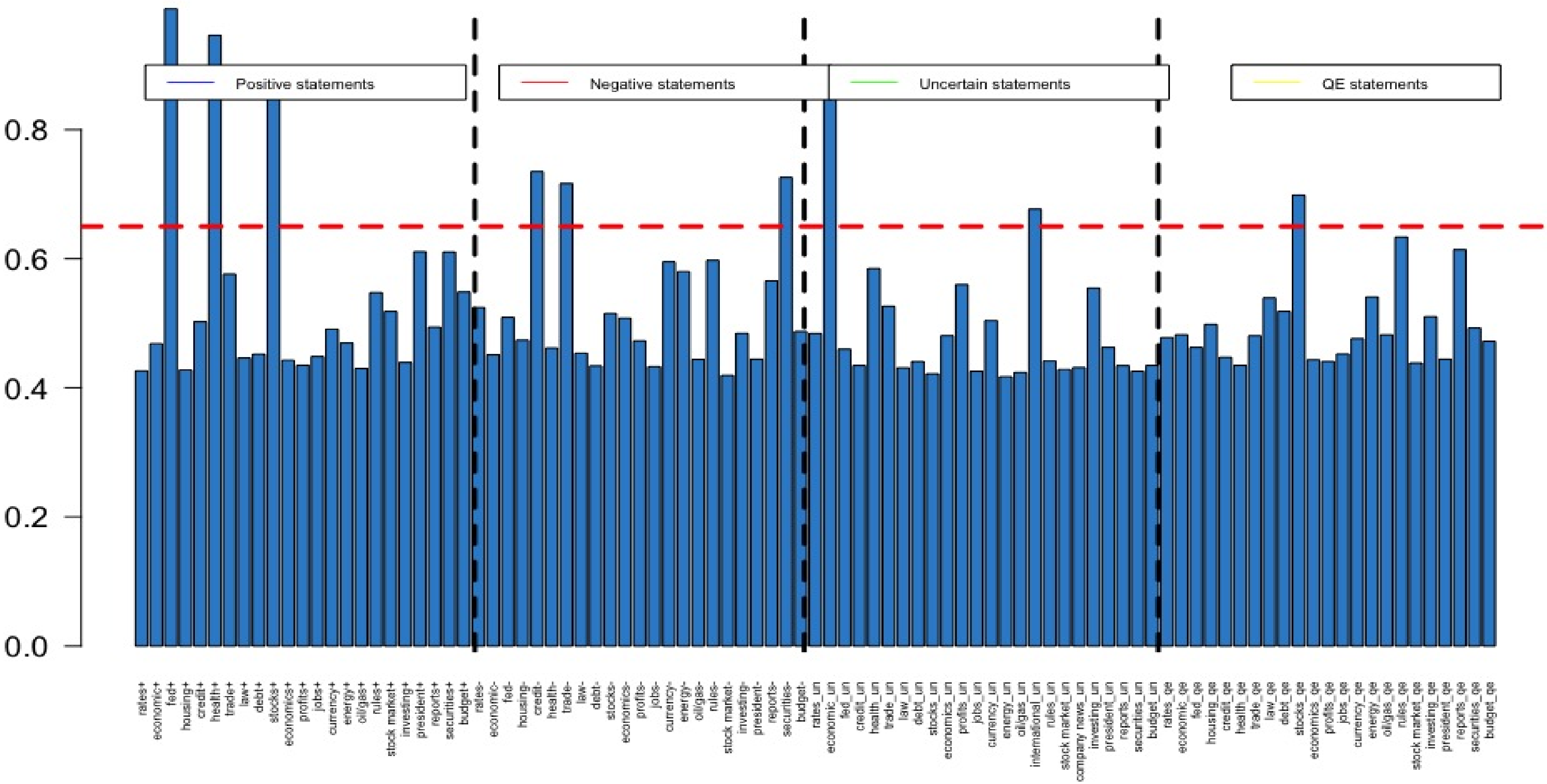


Figure E.5: Bayesian Lasso for surprises in S&P with asymmetric effect

Table E.1: Predictability of topic time series

	Dependent variable:									
	rates	computers	economic	food	people	media	fed	housing	credit	cars
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ NEER USA	−0.015	0.007	−0.003	−0.0002	−0.036*	0.006	−0.005	0.025	0.018	0.010
	(0.009)	(0.012)	(0.011)	(0.005)	(0.019)	(0.006)	(0.014)	(0.017)	(0.012)	(0.007)
Δ NEER Euro	0.020*	−0.012	0.013	0.002	0.037**	−0.005	0.030*	−0.022	−0.019*	−0.010
	(0.011)	(0.014)	(0.011)	(0.005)	(0.018)	(0.009)	(0.016)	(0.020)	(0.011)	(0.007)
Δ TEDRATE	−0.099	0.227	0.675**	−0.023	−0.373	0.119	−0.029	0.723***	0.190	0.168
	(0.179)	(0.260)	(0.286)	(0.080)	(0.260)	(0.150)	(0.307)	(0.154)	(0.255)	(0.105)
Δ S&P500	−0.0003*	0.0001	0.0003	−0.00002	0.0003	0.0001	−0.001**	0.0003	−0.0002	0.0002
	(0.0002)	(0.0002)	(0.0003)	(0.0001)	(0.0003)	(0.0001)	(0.0003)	(0.0002)	(0.0003)	(0.0003)
Δ AAA10Y	−0.004	−0.707	0.228	0.164	−0.055	0.538	−1.804**	−0.129	0.675	0.426
	(0.473)	(0.726)	(0.498)	(0.251)	(0.707)	(0.558)	(0.812)	(0.413)	(0.532)	(0.587)
Δ BAA10Y	0.132	0.371	−0.797*	−0.127	0.067	−0.481	0.758	−0.607*	−0.514	−0.404
	(0.346)	(0.457)	(0.413)	(0.256)	(0.542)	(0.593)	(0.700)	(0.313)	(0.612)	(0.518)
Constant	0.004	−0.002	0.001	0.001	−0.002	0.001	0.009	−0.001	0.002	−0.001
	(0.086)	(0.076)	(0.090)	(0.094)	(0.067)	(0.067)	(0.061)	(0.035)	(0.082)	(0.073)
Observations	195	195	195	195	195	195	195	195	195	195
R^2	0.016	0.019	0.063	0.001	0.070	0.008	0.113	0.068	0.024	0.011
Adjusted R^2	-0.016	-0.012	0.033	-0.031	0.041	-0.024	0.085	0.038	-0.007	-0.020

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table E.2: Predictability of topic time series

	Dependent variable:									
	health	trade	law	debt	loans	stocks	schools	economics	retailers	industry
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ NEER USA	0.005	−0.015	−0.010	0.011	−0.003	0.007	0.008	−0.007	−0.002	0.002
	(0.006)	(0.012)	(0.007)	(0.012)	(0.010)	(0.014)	(0.006)	(0.009)	(0.007)	(0.008)
Δ NEER Euro	−0.011	0.019	0.011	0.007	0.020	−0.001	−0.005	0.011	0.012	−0.005
	(0.008)	(0.015)	(0.008)	(0.015)	(0.013)	(0.016)	(0.007)	(0.010)	(0.009)	(0.009)
Δ TEDRATE	0.388	0.240	−0.189*	0.038	−0.044	0.148	−0.161	−0.152	0.076	0.843***
	(0.282)	(0.348)	(0.108)	(0.431)	(0.260)	(0.343)	(0.167)	(0.324)	(0.131)	(0.319)
Δ S&P500	−0.001	0.001	−0.00003	0.0005	−0.0003	0.0005**	0.0001	0.001**	0.0001	0.0002
	(0.0004)	(0.001)	(0.0001)	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0004)	(0.0002)	(0.0002)
Δ AAA10Y	−0.555	0.078	−0.008	−0.099	0.527	0.326	−0.060	−0.454	0.713	−0.079
	(0.574)	(0.640)	(0.481)	(0.705)	(0.649)	(0.467)	(0.418)	(0.430)	(0.672)	(0.768)
Δ BAA10Y	0.731	−0.288	0.148	−0.922	−0.796	−0.732**	−0.149	−0.128	−0.819	−0.212
	(0.581)	(0.393)	(0.327)	(0.643)	(0.565)	(0.366)	(0.369)	(0.292)	(0.709)	(0.513)
Constant	0.002	−0.005	0.0001	0.001	0.008	−0.002	−0.0001	−0.007	0.003	−0.003
	(0.079)	(0.087)	(0.097)	(0.070)	(0.097)	(0.084)	(0.073)	(0.075)	(0.072)	(0.088)
Observations	195	195	195	195	195	195	195	195	195	195
R^2	0.054	0.049	0.007	0.094	0.045	0.031	0.009	0.053	0.023	0.056
Adjusted R^2	0.024	0.019	-0.025	0.065	0.015	-0.0001	-0.022	0.023	-0.008	0.026

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table E.3: Predictability of topic time series

	Dependent variable:									
	cities	profits	jobs	currency	airlines	military	energy	oil/gas	international	hotels
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ NEER USA	0.003	0.005	0.004	−0.001	−0.005	−0.005	−0.007	−0.008	−0.012	0.025
	(0.008)	(0.010)	(0.009)	(0.010)	(0.010)	(0.015)	(0.009)	(0.010)	(0.008)	(0.017)
Δ NEER Euro	−0.001	0.00001	0.005	−0.005	0.014	0.003	0.002	0.011	0.023**	−0.013
	(0.008)	(0.011)	(0.010)	(0.013)	(0.011)	(0.013)	(0.010)	(0.012)	(0.010)	(0.014)
Δ TEDRATE	−0.156	0.261	0.334	0.036	−0.183	−0.048	−0.274**	0.266	−0.480	−0.010
	(0.183)	(0.185)	(0.203)	(0.509)	(0.173)	(0.164)	(0.136)	(0.305)	(0.444)	(0.139)
Δ S&P500	0.0001	0.0003	0.00001	0.00000	−0.00003	0.001*	0.001	0.0002	0.001	0.0001
	(0.0001)	(0.0002)	(0.0004)	(0.0003)	(0.0002)	(0.001)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
Δ AAA10Y	−0.709	−0.136	1.044**	0.048	−1.023	−0.961*	−0.349	−0.936*	−0.663	0.863
	(0.451)	(0.499)	(0.500)	(0.740)	(0.886)	(0.512)	(0.399)	(0.562)	(0.529)	(0.676)
Δ BAA10Y	0.348	−0.368	−1.159***	0.150	0.430	0.246	0.279	0.288	0.055	−0.993
	(0.361)	(0.315)	(0.418)	(0.886)	(0.330)	(0.286)	(0.343)	(0.468)	(0.293)	(0.710)
Constant	−0.001	−0.001	0.004	−0.001	0.001	−0.010	−0.008	−0.002	−0.003	0.003
	(0.076)	(0.086)	(0.085)	(0.093)	(0.070)	(0.071)	(0.086)	(0.087)	(0.081)	(0.069)
Observations	195	195	195	195	195	195	195	195	195	195
R^2	0.011	0.019	0.044	0.005	0.030	0.061	0.040	0.034	0.068	0.057
Adjusted R^2	-0.021	-0.012	0.013	-0.027	-0.001	0.031	0.009	0.003	0.038	0.027
Note:	HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01									

Table E.4: Predictability of topic time series

	Dependent variable:									
	rules	stock market	company news	services	investing	president	reports	securities	budget	deals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ NEER USA	−0.005	0.001	−0.022**	0.00002	0.005	0.012	0.001	0.001	−0.005	−0.002
	(0.011)	(0.008)	(0.009)	(0.006)	(0.010)	(0.011)	(0.010)	(0.009)	(0.010)	(0.006)
Δ NEER Euro	0.024**	0.016*	0.032***	−0.008	0.010	−0.018	0.006	0.007	0.012	0.004
	(0.011)	(0.009)	(0.012)	(0.010)	(0.012)	(0.013)	(0.011)	(0.009)	(0.012)	(0.009)
Δ TEDRATE	0.171	0.238	0.003	−0.389	0.534*	0.759***	0.533**	−0.391	0.350*	0.138
	(0.262)	(0.253)	(0.206)	(0.269)	(0.303)	(0.229)	(0.258)	(0.329)	(0.183)	(0.205)
Δ S&P500	−0.001**	0.0003	−0.0002	0.002*	0.0001	−0.0002	0.0004*	0.0001	0.001***	−0.0003
	(0.0005)	(0.0004)	(0.0002)	(0.001)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Δ AAA10Y	−0.227	0.551	0.444	−0.056	−0.464	0.056	1.240**	−1.215**	1.076*	1.586**
	(0.557)	(0.700)	(0.710)	(0.654)	(0.737)	(0.415)	(0.562)	(0.529)	(0.609)	(0.739)
Δ BAA10Y	−0.018	−1.147**	−0.338	−0.299	−0.425	−0.116	−1.393***	0.509	−1.233**	−0.976*
	(0.397)	(0.574)	(0.665)	(0.416)	(0.609)	(0.379)	(0.455)	(0.434)	(0.517)	(0.588)
Constant	0.015	0.002	0.004	−0.019	0.002	0.001	0.001	−0.0003	−0.002	0.005
	(0.043)	(0.082)	(0.089)	(0.069)	(0.091)	(0.095)	(0.089)	(0.055)	(0.099)	(0.083)
Observations	195	195	195	195	195	195	195	195	195	195
R^2	0.098	0.066	0.033	0.208	0.055	0.059	0.068	0.047	0.062	0.042
Adjusted R^2	0.069	0.037	0.002	0.182	0.025	0.029	0.039	0.016	0.032	0.012

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Appendix F. Effect of Central bank communication

F.1 Central bank communication and Yield curves

Table F.1: ΔYields, one day difference

	Dependent variable:							
	1 Year	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Economic aggregated	−0.114	0.741	1.672*	1.818**	1.720**	1.608**	1.486**	1.357**
	(0.810)	(0.848)	(0.930)	(0.908)	(0.795)	(0.701)	(0.666)	(0.677)
Fed news	−0.193	0.017	−0.172	−0.350	−0.319	−0.271	−0.250	−0.264
	(0.910)	(1.123)	(1.146)	(1.015)	(0.896)	(0.833)	(0.781)	(0.753)
Fed news on FG dates	0.312	−0.506	−1.214	−1.110	−0.806	−0.654	−0.655	−0.742
	(0.805)	(0.855)	(0.981)	(0.911)	(0.820)	(0.759)	(0.742)	(0.752)
Constant	−0.003	−0.008	−0.013	−0.014	−0.013	−0.012	−0.011	−0.009
	(0.011)	(0.014)	(0.014)	(0.014)	(0.013)	(0.012)	(0.011)	(0.010)
Observations	186	186	186	186	186	186	186	186
R^2	0.001	0.006	0.023	0.028	0.030	0.031	0.030	0.027
Adjusted R^2	-0.015	-0.011	0.007	0.012	0.014	0.016	0.014	0.011
Note:	HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01							

Table F.2: Δ Yields, one day difference

	Dependent variable:							
	1 Year	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	0.721	1.459	3.110*	3.739**	3.329**	2.838***	2.471***	2.231**
	(0.935)	(1.226)	(1.723)	(1.830)	(1.460)	(1.086)	(0.909)	(0.936)
ΔFed news	0.496	0.530	−1.144	−1.767	−1.090	−0.600	−0.591	−0.927
	(0.997)	(1.320)	(1.628)	(1.420)	(1.079)	(0.921)	(0.974)	(1.075)
Δ Fed news on FG dates	−0.066	−0.743	0.040	0.564	0.179	−0.080	0.135	0.739
	(1.753)	(2.161)	(2.478)	(2.145)	(1.676)	(1.509)	(1.561)	(1.636)
Constant	−0.005	−0.003	−0.004	−0.004	−0.002	−0.001	−0.0003	−0.001
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
Observations	185	185	185	185	185	185	185	185
R^2	0.010	0.021	0.048	0.079	0.088	0.082	0.069	0.053
Adjusted R^2	-0.006	0.004	0.032	0.064	0.073	0.067	0.053	0.038

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.3: Δ Forward Rates, one day difference

	Dependent variable:							
	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year	1 Year Forward 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Economic aggregated	2.036**	2.290**	1.678**	1.396*	1.144	0.855	0.581	2.371**
	(1.034)	(1.125)	(0.852)	(0.745)	(0.717)	(0.869)	(1.223)	(1.145)
Fed news	0.088	−0.556	−0.390	−0.163	−0.122	−0.237	−0.448	−0.501
	(1.366)	(1.242)	(0.924)	(0.915)	(0.876)	(1.050)	(1.517)	(1.306)
Fed news on FG dates	−1.602	−1.559	−0.472	−0.079	−0.389	−0.924	−1.414	−1.660
	(1.118)	(1.099)	(0.844)	(0.785)	(0.770)	(0.918)	(1.250)	(1.136)
Constant	−0.015	−0.016	−0.013	−0.011	−0.008	−0.003	0.002	−0.017
	(0.017)	(0.016)	(0.014)	(0.013)	(0.011)	(0.011)	(0.014)	(0.017)
Observations	186	186	186	186	186	186	186	186
R^2	0.028	0.030	0.023	0.021	0.014	0.012	0.013	0.031
Adjusted R^2	0.012	0.014	0.007	0.005	-0.002	-0.004	-0.003	0.015

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.4: Δ Forward Rates, one day difference

	Dependent variable:							
	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year	1 Year Forward 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	2.855	4.940**	3.417**	1.772***	1.089	0.988	1.075	4.913**
	(1.805)	(2.421)	(1.334)	(0.671)	(0.735)	(0.982)	(1.896)	(2.467)
Δ Fed news	−0.198	−3.323*	−0.925	1.009	0.374	−1.582	−3.575*	−3.252*
	(1.890)	(1.835)	(1.157)	(1.683)	(1.940)	(1.785)	(2.040)	(1.862)
Δ Fed news on FG dates	−0.973	1.427	0.272	−1.140	−0.196	2.355	5.110*	1.328
	(2.836)	(2.758)	(1.519)	(2.101)	(2.490)	(2.436)	(2.735)	(2.851)
Constant	−0.001	−0.006	−0.001	0.004	0.002	−0.001	−0.005	−0.006
	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)	(0.006)	(0.008)	(0.007)
Observations	185	185	185	185	185	185	185	185
R^2	0.035	0.082	0.087	0.045	0.016	0.010	0.016	0.075
Adjusted R^2	0.019	0.067	0.071	0.029	-0.001	-0.006	0.0001	0.060

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.5: Δ Yields, two days difference

	Dependent variable:							
	1 Year	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Economic aggregated	0.171	1.476	3.046***	3.276***	2.883**	2.414**	1.913**	1.385
	(0.981)	(0.950)	(1.147)	(1.231)	(1.174)	(1.064)	(0.971)	(0.949)
Fed news	0.377	1.107	1.722	2.349*	2.472*	2.167*	1.622	0.969
	(1.116)	(1.286)	(1.427)	(1.404)	(1.298)	(1.226)	(1.200)	(1.275)
Fed news on FG dates	−0.109	−1.343	−2.293*	−2.454**	−2.277*	−1.969*	−1.589	−1.167
	(0.824)	(0.901)	(1.186)	(1.222)	(1.173)	(1.130)	(1.110)	(1.131)
Constant	−0.020	−0.031*	−0.051***	−0.062***	−0.058***	−0.049***	−0.038**	−0.025
	(0.014)	(0.016)	(0.017)	(0.018)	(0.017)	(0.016)	(0.015)	(0.016)
Observations	186	186	186	186	186	186	186	186
R^2	0.001	0.027	0.066	0.075	0.074	0.063	0.043	0.020
Adjusted R^2	-0.015	0.011	0.051	0.060	0.058	0.048	0.027	0.003

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.6: Δ Yields, two days difference

	Dependent variable:							
	1 Year	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	−0.702	0.229	2.777	4.250*	3.962*	3.257*	2.566*	1.999*
	(0.991)	(1.203)	(2.032)	(2.451)	(2.164)	(1.701)	(1.322)	(1.205)
Δ Fed news	1.676	2.413	1.442	1.029	1.716	2.001	1.561	0.634
	(1.269)	(1.472)	(1.360)	(1.761)	(1.862)	(1.879)	(2.126)	(2.480)
Δ Fed news on FG dates	−2.183	−2.862	−2.039	−1.288	−1.621	−1.696	−1.005	0.279
	(1.380)	(1.806)	(2.241)	(2.621)	(2.514)	(2.461)	(2.681)	(3.124)
Constant	−0.016***	−0.010	−0.011	−0.012*	−0.011	−0.009	−0.007	−0.006
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)
Observations	185	185	185	185	185	185	185	185
R^2	0.007	0.013	0.041	0.073	0.081	0.073	0.052	0.028
Adjusted R^2	-0.009	-0.003	0.025	0.058	0.066	0.058	0.037	0.012

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.7: Δ Forward Rates, two days difference

	Dependent variable:							
	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year	1 Year Forward 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Economic aggregated	3.432***	4.225***	2.720*	1.536	0.476	−0.673	−1.817	4.355***
	(1.298)	(1.514)	(1.398)	(1.115)	(1.125)	(1.544)	(2.144)	(1.523)
Fed news	1.950	2.465	3.154**	2.092	0.359	−1.460	−3.089	2.371
	(1.636)	(1.754)	(1.401)	(1.305)	(1.473)	(2.177)	(3.264)	(1.805)
Fed news on FG dates	−2.873**	−2.857*	−2.313*	−1.506	−0.569	0.442	1.434	−2.945*
	(1.292)	(1.524)	(1.340)	(1.218)	(1.246)	(1.563)	(2.119)	(1.548)
Constant	−0.049***	−0.074***	−0.065***	−0.037**	−0.007	0.022	0.051	−0.074***
	(0.019)	(0.022)	(0.020)	(0.017)	(0.018)	(0.030)	(0.048)	(0.022)
Observations	186	186	186	186	186	186	186	186
R^2	0.072	0.071	0.060	0.032	0.002	0.006	0.014	0.073
Adjusted R^2	0.056	0.056	0.044	0.016	-0.014	-0.011	-0.002	0.057

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.8: Δ Forward Rates, two days difference

	Dependent variable:							
	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year	1 Year Forward 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	2.151	5.946*	4.739**	2.117*	0.330	−0.616	−0.968	5.732*
	(1.919)	(3.076)	(2.350)	(1.244)	(1.355)	(1.992)	(2.697)	(3.060)
Δ Fed news	2.603	−0.220	2.030	3.578	1.628	−2.128	−5.756	−0.059
	(1.812)	(2.085)	(2.576)	(2.748)	(3.223)	(4.198)	(5.505)	(1.973)
Δ Fed news on FG dates	−3.058	−0.426	−1.362	−2.744	−0.491	4.205	9.105	−0.673
	(2.575)	(3.287)	(3.021)	(3.246)	(4.041)	(5.250)	(6.789)	(3.229)
Constant	−0.005	−0.015	−0.011	−0.004	−0.001	0.001	0.003	−0.015
	(0.008)	(0.009)	(0.008)	(0.007)	(0.008)	(0.011)	(0.016)	(0.009)
Observations	185	185	185	185	185	185	185	185
R^2	0.031	0.077	0.086	0.054	0.007	0.007	0.016	0.072
Adjusted R^2	0.015	0.062	0.071	0.038	-0.010	-0.009	0.0002	0.056

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.9: Δ TIPS Yields, one day difference

	Dependent variable:							
	5 Year	10 Year	15 Year	20 Year	5 Year F	10 Year F	15 Year F	20 Year F
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Economic aggregated	2.526^{**}	2.220^{*}	2.011^{*}	1.952^{**}	2.568^{**}	1.538	1.729^{*}	1.678^{**}
	(1.264)	(1.163)	(1.066)	(0.958)	(1.305)	(1.142)	(0.980)	(0.841)
Fed news	−1.006	−1.004	−0.922	−0.781	−1.116	−0.930	−0.525	−0.346
	(1.688)	(1.378)	(1.204)	(1.105)	(1.465)	(1.202)	(1.094)	(1.135)
Fed news on FG dates	-1.879^{*}	−1.208	−1.014	−1.027	−1.318	−0.323	−0.951	−0.998
	(1.139)	(1.027)	(0.930)	(0.853)	(1.124)	(1.012)	(0.895)	(0.804)
Constant	−0.008	−0.010	−0.011	−0.011	−0.012	−0.010	−0.014	−0.011
	(0.026)	(0.021)	(0.019)	(0.018)	(0.021)	(0.021)	(0.018)	(0.018)
Observations	132	132	132	132	132	132	132	132
R^2	0.041	0.034	0.036	0.045	0.036	0.017	0.040	0.046
Adjusted R^2	0.019	0.011	0.014	0.022	0.013	-0.006	0.018	0.024

Note: HAC standard errors are in parentheses; $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table F.10: Δ TIPS Yields, one day difference

	Dependent variable:							
	5 Year	10 Year	15 Year	20 Year	5 Year F	10 Year F	15 Year F	20 Year F
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	4.813**	4.587**	4.324**	3.982**	4.856*	4.013*	3.530***	2.155***
	(2.372)	(2.263)	(2.067)	(1.749)	(2.588)	(2.066)	(1.280)	(0.810)
Δ Fed news	−3.915	−3.428	−2.579	−1.709	−4.251*	−1.830	0.086	1.519
	(2.579)	(2.243)	(1.735)	(1.239)	(2.534)	(1.607)	(0.889)	(1.769)
Δ Fed news on FG dates	1.206	1.549	0.909	0.009	2.361	0.959	−1.713	−3.351
	(3.510)	(2.971)	(2.342)	(1.810)	(3.243)	(2.228)	(1.356)	(2.126)
Constant	−0.005	−0.005	−0.006	−0.005	−0.006	−0.006	−0.006	−0.002
	(0.009)	(0.008)	(0.006)	(0.006)	(0.008)	(0.006)	(0.004)	(0.005)
Observations	132	132	132	132	132	132	132	132
R^2	0.082	0.099	0.123	0.140	0.092	0.113	0.150	0.102
Adjusted R^2	0.060	0.078	0.102	0.120	0.071	0.092	0.130	0.081

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.11: Δ Inflation Compensation, one day difference

	Dependent variable:				
	2 Year	5 Year	10 Year	15 Year	20 Year
	(1)	(2)	(3)	(4)	(5)
Economic aggregated	−0.279	−0.504	−0.311	−0.488	−0.728*
	(0.727)	(0.434)	(0.374)	(0.363)	(0.387)
Fed news	−0.832	0.304	0.189	0.004	−0.158
	(1.177)	(0.593)	(0.446)	(0.451)	(0.490)
Fed news on FG dates	0.617	0.252	−0.265	−0.004	0.286
	(0.611)	(0.435)	(0.354)	(0.332)	(0.355)
Constant	0.018	0.006	0.008	0.013*	0.017**
	(0.018)	(0.008)	(0.006)	(0.007)	(0.008)
Observations	132	132	132	132	132
R^2	0.005	0.008	0.007	0.010	0.020
Adjusted R^2	-0.018	-0.015	-0.016	-0.013	-0.003

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.12: Δ Inflation Compensation, one day difference

	Dependent variable:				
	2 Year	5 Year	10 Year	15 Year	20 Year
	(1)	(2)	(3)	(4)	(5)
ΔEconomic aggregated	-2.301^{*}	-0.945	-0.158	-0.520	-0.814
	(1.174)	(0.633)	(0.411)	(0.531)	(0.650)
ΔFed news	2.170	2.495^{**}	1.301	1.243	0.974
	(1.886)	(1.185)	(0.823)	(0.858)	(1.018)
Δ Fed news on FG dates	-0.782	-1.197	-0.860	-0.598	0.001
	(2.526)	(1.472)	(1.040)	(1.106)	(1.336)
Constant	0.008	0.006	0.005	0.007^{**}	0.008^{**}
	(0.007)	(0.004)	(0.003)	(0.003)	(0.004)
Observations	132	132	132	132	132
R^2	0.040	0.051	0.019	0.020	0.025
Adjusted R^2	0.017	0.029	-0.004	-0.003	0.002

Note: HAC standard errors are in parentheses; $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table F.13: Δ TIPS Yields, two days difference

	Dependent variable:							
	5 Year	10 Year	15 Year	20 Year	5 Year F	10 Year F	15 Year F	20 Year F
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Economic aggregated	3.517**	3.159**	2.611**	2.504**	4.096**	1.771	1.608*	2.909**
	(1.443)	(1.394)	(1.195)	(1.069)	(1.794)	(1.224)	(0.945)	(1.195)
Fed news	2.244	2.010	1.967	2.134	2.065	1.546	2.366	2.528*
	(2.323)	(1.909)	(1.656)	(1.539)	(2.051)	(1.584)	(1.508)	(1.355)
Fed news on FG dates	−2.459*	−1.926	−1.523	−1.555	−2.352	−0.730	−0.974	−2.493**
	(1.351)	(1.289)	(1.190)	(1.143)	(1.518)	(1.284)	(1.186)	(1.214)
Constant	−0.072*	−0.066**	−0.062**	−0.063***	−0.073**	−0.050**	−0.063***	−0.063***
	(0.038)	(0.031)	(0.026)	(0.023)	(0.031)	(0.024)	(0.019)	(0.019)
Observations	132	132	132	132	132	132	132	132
R^2	0.048	0.050	0.050	0.061	0.057	0.027	0.051	0.094
Adjusted R^2	0.026	0.027	0.028	0.039	0.035	0.004	0.029	0.073

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.14: Δ TIPS Yields, two days difference

	Dependent variable:							
	5 Year	10 Year	15 Year	20 Year	5 Year F	10 Year F	15 Year F	20 Year F
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	5.579**	5.433**	4.605*	4.271**	6.660**	3.719	2.710***	4.068***
	(2.828)	(2.645)	(2.472)	(2.071)	(3.130)	(2.337)	(0.921)	(1.395)
ΔFed news	−2.209	−2.501	−1.673	−0.576	−4.136	−1.497	1.507	3.505*
	(3.118)	(2.732)	(2.082)	(1.513)	(3.458)	(1.967)	(1.606)	(2.012)
Δ Fed news on FG dates	−1.269	−0.034	−0.133	−0.878	1.448	0.687	−1.583	−4.662*
	(4.067)	(3.537)	(2.722)	(2.142)	(4.231)	(2.656)	(2.246)	(2.405)
Constant	−0.018	−0.015	−0.015*	−0.016**	−0.013	−0.013	−0.019***	−0.013*
	(0.012)	(0.010)	(0.009)	(0.008)	(0.012)	(0.009)	(0.007)	(0.007)
Observations	132	132	132	132	132	132	132	132
R^2	0.068	0.082	0.082	0.093	0.088	0.055	0.069	0.153
Adjusted R^2	0.047	0.060	0.060	0.072	0.067	0.033	0.047	0.133

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.15: Δ Inflation Compensation, two days difference

	Dependent variable:				
	2 Year	5 Year	10 Year	15 Year	20 Year
	(1)	(2)	(3)	(4)	(5)
Economic aggregated	−0.446	−0.281	0.165	0.031	−0.519
	(0.962)	(0.704)	(0.678)	(0.721)	(0.704)
Fed news	−1.766	0.394	1.177	1.082	0.551
	(2.603)	(1.406)	(0.827)	(0.827)	(0.907)
Fed news on FG dates	0.136	−0.327	−0.857	−0.933	−0.463
	(0.975)	(0.751)	(0.646)	(0.650)	(0.685)
Constant	0.041	0.012	−0.005	0.001	0.012
	(0.046)	(0.026)	(0.013)	(0.012)	(0.013)
Observations	132	132	132	132	132
R^2	0.008	0.001	0.011	0.012	0.005
Adjusted R^2	-0.016	-0.022	-0.012	-0.011	-0.018

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.16: Δ Inflation Compensation, two days difference

	Dependent variable:				
	2 Year	5 Year	10 Year	15 Year	20 Year
	(1)	(2)	(3)	(4)	(5)
ΔEconomic aggregated	-3.740^{**}	-2.018	-0.351	0.034	-0.500
	(1.862)	(1.344)	(0.773)	(0.742)	(0.876)
Δ Fed news	-0.0005	2.973	2.132	2.058	1.544
	(3.716)	(2.627)	(1.956)	(1.720)	(1.978)
Δ Fed news on FG dates	2.071	-0.784	-0.677	-0.806	-0.316
	(4.367)	(3.117)	(2.364)	(2.097)	(2.454)
Constant	0.013	0.010	0.005	0.007	0.009^{*}
	(0.014)	(0.009)	(0.006)	(0.005)	(0.005)
Observations	132	132	132	132	132
R^2	0.039	0.030	0.021	0.025	0.012
Adjusted R^2	0.016	0.007	-0.002	0.002	-0.011

Note: HAC standard errors are in parentheses; $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

F.2 Central bank communication and MP shocks

Table F.17: Δ Yields, one day difference

	Dependent variable:							
	1 Year	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	1.401	2.373	4.138*	4.627*	3.902**	3.128**	2.554**	2.166*
	(0.939)	(1.458)	(2.463)	(2.583)	(1.961)	(1.415)	(1.168)	(1.168)
Δ Fed news	1.013	1.121	−0.751	−1.739	−0.935	−0.242	−0.113	−0.409
	(1.672)	(2.018)	(2.485)	(2.390)	(1.704)	(1.295)	(1.326)	(1.529)
ΔFed news on FG dates	−1.360	−2.571	−1.906	−0.596	−0.774	−1.134	−0.936	−0.162
	(1.638)	(2.110)	(2.790)	(2.721)	(1.956)	(1.565)	(1.689)	(2.045)
PCA_ffr4_hf	0.478***	0.487***	0.373***	0.211**	0.125*	0.075	0.042	0.015
	(0.105)	(0.117)	(0.111)	(0.086)	(0.076)	(0.069)	(0.062)	(0.059)
Constant	−0.009***	−0.007	−0.008	−0.006	−0.003	−0.001	0.0002	0.0003
	(0.003)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Observations	155	155	155	155	155	155	155	155
R^2	0.382	0.308	0.180	0.131	0.115	0.095	0.071	0.048
Adjusted R^2	0.365	0.289	0.158	0.108	0.091	0.070	0.046	0.023

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.18: Δ Yields, two days difference

	Dependent variable:							
	1 Year	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	0.055	1.281	4.255*	5.706*	5.006**	3.912**	2.919**	2.082
	(0.972)	(1.399)	(2.395)	(2.946)	(2.532)	(1.924)	(1.469)	(1.457)
Δ Fed news	1.570	2.348	1.146	0.502	1.482	2.038	1.658	0.609
	(1.308)	(1.743)	(2.200)	(2.586)	(2.641)	(2.648)	(2.815)	(3.002)
ΔFed news on FG dates	−2.476*	−3.875**	−3.073	−1.560	−1.762	−1.911	−1.084	0.650
	(1.372)	(1.827)	(2.735)	(3.172)	(3.024)	(3.146)	(3.650)	(4.011)
PCA_ffr4_hf	0.607***	0.578***	0.470***	0.310**	0.219	0.158	0.102	0.042
	(0.083)	(0.116)	(0.146)	(0.146)	(0.147)	(0.135)	(0.117)	(0.121)
Constant	−0.016***	−0.011*	−0.013	−0.014	−0.012	−0.009	−0.006	−0.003
	(0.005)	(0.007)	(0.009)	(0.010)	(0.009)	(0.008)	(0.008)	(0.008)
Observations	155	155	155	155	155	155	155	155
R^2	0.420	0.299	0.170	0.129	0.115	0.095	0.062	0.029
Adjusted R^2	0.405	0.281	0.148	0.105	0.091	0.071	0.037	0.003

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.19: Δ Forward Rates, one day difference

	Dependent variable:							
	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year	1 Year Forward 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	4.044	5.984*	3.753**	1.392	0.402	0.208	0.233	5.997*
	(2.477)	(3.421)	(1.730)	(0.937)	(0.979)	(1.188)	(2.245)	(3.491)
Δ Fed news	0.458	−3.538	−0.991	1.814	1.406	−0.738	−2.960	−3.313
	(2.678)	(3.090)	(1.733)	(2.060)	(2.404)	(2.208)	(3.142)	(3.162)
ΔFed news on FG dates	−3.462	0.155	0.197	−2.210	−1.633	1.663	5.747	−0.259
	(3.056)	(3.738)	(2.015)	(2.535)	(2.968)	(3.006)	(4.281)	(3.800)
PCA_ffr4_hf	0.454***	0.163*	−0.022	−0.064	−0.082	−0.105	−0.136	0.206**
	(0.140)	(0.084)	(0.068)	(0.070)	(0.068)	(0.077)	(0.092)	(0.093)
Constant	−0.006	−0.009	−0.0005	0.006	0.006	0.003	−0.001	−0.009
	(0.007)	(0.008)	(0.007)	(0.005)	(0.005)	(0.006)	(0.009)	(0.008)
Observations	155	155	155	155	155	155	155	155
R^2	0.188	0.103	0.089	0.050	0.023	0.016	0.027	0.105
Adjusted R^2	0.167	0.079	0.065	0.024	-0.003	-0.010	0.001	0.081

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.20: Δ Forward Rates, two days difference

	Dependent variable:							
	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year	1 Year Forward 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Economic aggregated	3.677	7.807**	5.547*	1.866	−0.374	−1.639	−2.518	7.624**
	(2.328)	(3.751)	(2.903)	(1.552)	(1.795)	(2.891)	(3.793)	(3.671)
Fed news	2.550	−1.084	1.776	4.380	2.342	−2.250	−6.896	−0.812
	(2.739)	(3.106)	(3.573)	(3.441)	(4.022)	(4.849)	(6.737)	(3.015)
Fed news on FG dates	−4.832	−0.629	−0.756	−3.098	−0.735	5.593	13.036	−1.162
	(3.139)	(4.175)	(3.913)	(4.025)	(5.146)	(6.480)	(8.800)	(4.106)
PCA_ffr4_hf	0.513***	0.280*	0.068	0.009	−0.065	−0.187	−0.339	0.325*
	(0.181)	(0.161)	(0.167)	(0.142)	(0.126)	(0.250)	(0.424)	(0.169)
Constant	−0.007	−0.017	−0.011	−0.003	0.003	0.007	0.012	−0.017
	(0.010)	(0.012)	(0.010)	(0.009)	(0.009)	(0.012)	(0.018)	(0.012)
Observations	155	155	155	155	155	155	155	155
R^2	0.162	0.113	0.095	0.055	0.008	0.021	0.042	0.116
Adjusted R^2	0.140	0.089	0.071	0.030	-0.018	-0.005	0.017	0.092

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.21: Δ TIPS Yields, one day difference

	Dependent variable:							
	5 Year	10 Year	15 Year	20 Year	5 Year F	10 Year F	15 Year F	20 Year F
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔEconomic aggregated	7.996**	7.076**	6.214**	5.435**	7.476**	5.147*	3.869**	2.173**
	(3.289)	(3.223)	(2.759)	(2.250)	(3.622)	(2.664)	(1.617)	(1.071)
ΔFed news	−5.656	−4.372	−3.051	−1.752	−5.029	−1.652	0.906	3.216
	(4.840)	(3.760)	(3.119)	(2.608)	(3.896)	(2.896)	(2.160)	(3.292)
Δ Fed news on FG dates	−0.274	0.004	−0.416	−1.370	0.397	−0.084	−2.727	−5.562
	(5.324)	(4.126)	(3.463)	(2.981)	(4.289)	(3.200)	(2.624)	(3.625)
PCA_ffr4_hf	−0.041	−0.092	−0.100	−0.100	−0.129	−0.137	−0.103	−0.098
	(0.330)	(0.289)	(0.251)	(0.220)	(0.330)	(0.220)	(0.160)	(0.126)
Constant	−0.006	−0.005	−0.004	−0.004	−0.005	−0.003	−0.004	−0.0001
	(0.012)	(0.010)	(0.009)	(0.008)	(0.012)	(0.008)	(0.007)	(0.006)
Observations	104	104	104	104	104	104	104	104
R^2	0.149	0.161	0.169	0.174	0.148	0.134	0.138	0.107
Adjusted R^2	0.115	0.127	0.136	0.140	0.114	0.099	0.103	0.071
Note:								*p<0.1; **p<0.05; ***p<0.01

Table F.22: Δ TIPS Yields, two days difference

	Dependent variable:							
	5 Year	10 Year	15 Year	20 Year	5 Year F	10 Year F	15 Year F	20 Year F
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EC	7.996**	7.076**	6.214**	5.435**	7.476**	5.147*	3.869**	2.173**
	(3.289)	(3.223)	(2.759)	(2.250)	(3.622)	(2.664)	(1.617)	(1.071)
X7	−5.656	−4.372	−3.051	−1.752	−5.029	−1.652	0.906	3.216
	(4.840)	(3.760)	(3.119)	(2.608)	(3.896)	(2.896)	(2.160)	(3.292)
FG	−0.274	0.004	−0.416	−1.370	0.397	−0.084	−2.727	−5.562
	(5.324)	(4.126)	(3.463)	(2.981)	(4.289)	(3.200)	(2.624)	(3.625)
pc1ff1_hf	−0.041	−0.092	−0.100	−0.100	−0.129	−0.137	−0.103	−0.098
	(0.330)	(0.289)	(0.251)	(0.220)	(0.330)	(0.220)	(0.160)	(0.126)
Constant	−0.006	−0.005	−0.004	−0.004	−0.005	−0.003	−0.004	−0.0001
	(0.012)	(0.010)	(0.009)	(0.008)	(0.012)	(0.008)	(0.007)	(0.006)
Observations	104	104	104	104	104	104	104	104
R^2	0.149	0.161	0.169	0.174	0.148	0.134	0.138	0.107
Adjusted R^2	0.115	0.127	0.136	0.140	0.114	0.099	0.103	0.071
Note:	HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01							

Table F.23: Δ Inflation Compensation, one day difference

	Dependent variable:				
	2 Year	5 Year	10 Year	15 Year	20 Year
	(1)	(2)	(3)	(4)	(5)
Δ Economic aggregated	−4.198**	−2.277**	−0.871*	−1.020	−1.303
	(1.740)	(0.985)	(0.501)	(0.749)	(0.924)
ΔFed news	6.007*	5.345**	2.942**	2.526*	2.174
	(3.073)	(2.329)	(1.190)	(1.293)	(1.600)
Δ Fed news on FG date	−1.977	−2.509	−1.736	−1.446	−0.839
	(2.911)	(2.047)	(1.359)	(1.525)	(1.900)
PCA_ffr4_hf	−0.011	0.041	0.031	0.046	0.104
	(0.116)	(0.096)	(0.060)	(0.064)	(0.076)
Constant	0.010	0.007	0.005	0.008**	0.009**
	(0.008)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	104	104	104	104	104
R^2	0.127	0.153	0.062	0.052	0.071
Adjusted R^2	0.091	0.119	0.024	0.014	0.033

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

Table F.24: Δ Inflation Compensation, two days difference

	Dependent variable:				
	2 Year	5 Year	10 Year	15 Year	20 Year
	(1)	(2)	(3)	(4)	(5)
ΔEconomic aggregated	−4.198**	−2.277**	−0.871*	−1.020	−1.303
	(1.740)	(0.985)	(0.501)	(0.749)	(0.924)
ΔFed news	6.007*	5.345**	2.942**	2.526*	2.174
	(3.073)	(2.329)	(1.190)	(1.293)	(1.600)
ΔFed news on FG dates	−1.977	−2.509	−1.736	−1.446	−0.839
	(2.911)	(2.047)	(1.359)	(1.525)	(1.900)
PCA_ffr4_hf	−0.011	0.041	0.031	0.046	0.104
	(0.116)	(0.096)	(0.060)	(0.064)	(0.076)
Constant	0.010	0.007	0.005	0.008**	0.009**
	(0.008)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	104	104	104	104	104
R^2	0.127	0.153	0.062	0.052	0.071
Adjusted R^2	0.091	0.119	0.024	0.014	0.033
Note:	HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01				

F.3 On measurement error and bad controls

Table F.25: Δ Yields, one day difference

	Dependent variable:							
	1 Year	2 Year	5 Year	10 Year	15 Year	20 Year	25 Year	30 Year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PCA_ffr4_hf	0.461***	0.460***	0.335***	0.172**	0.090	0.045	0.017	−0.005
	(0.099)	(0.109)	(0.105)	(0.085)	(0.075)	(0.068)	(0.061)	(0.057)
Constant	−0.009***	−0.007	−0.007	−0.006	−0.003	−0.0002	0.001	0.001
	(0.003)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Observations	155	155	155	155	155	155	155	155
R^2	0.350	0.254	0.101	0.030	0.011	0.003	0.001	0.00005
Adjusted R^2	0.346	0.249	0.095	0.024	0.005	-0.003	-0.006	-0.006

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.26: Δ TIPS Yields, one day difference

	Dependent variable:							
	5 Year	10 Year	15 Year	20 Year	5 Year F	10 Year F	15 Year F	20 Year F
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PCA_ffr4_hf	−0.085	−0.127	−0.133	−0.135	−0.163	−0.159	−0.138	−0.146
	(0.349)	(0.305)	(0.264)	(0.231)	(0.344)	(0.228)	(0.170)	(0.135)
Constant	−0.005	−0.004	−0.003	−0.003	−0.004	−0.002	−0.003	0.001
	(0.013)	(0.011)	(0.009)	(0.008)	(0.011)	(0.008)	(0.006)	(0.006)
Observations	104	104	104	104	104	104	104	104
R^2	0.002	0.007	0.010	0.013	0.009	0.014	0.016	0.020
Adjusted R^2	-0.008	-0.003	0.0001	0.003	-0.001	0.005	0.006	0.010

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.27: Δ Inflation Compensation, one day difference

	Dependent variable:				
	2 Year	5 Year	10 Year	15 Year	20 Year
	(1)	(2)	(3)	(4)	(5)
PCA_ffr4_hf	0.001	0.039	0.025	0.043	0.106
	(0.130)	(0.106)	(0.065)	(0.065)	(0.074)
Constant	0.009	0.007	0.005	0.008**	0.009**
	(0.008)	(0.005)	(0.004)	(0.004)	(0.004)
Observations	104	104	104	104	104
R^2	0.00000	0.003	0.002	0.006	0.030
Adjusted R^2	-0.010	-0.006	-0.008	-0.004	0.020

Note: HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01

Table F.28: Δ Yields, one day difference

	Dependent variable:	
	1 Year	high frequency surprises
	(1)	(2)
PCA_ffr4_hf	0.481***	
	(0.137)	
PCA_S&P500_hf	0.002	
	(0.011)	
ff4_hf	−0.021	
	(0.268)	
Economic aggregated		−1.655
		(1.485)
Fed news on FG dates		1.154
		(2.720)
Fed news		−0.085
		(2.877)
Constant	−0.009**	0.002
	(0.005)	(0.007)
Observations	155	155
R^2	0.350	0.017
Adjusted R^2	0.337	-0.003
Note:	HAC standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01	

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):

$$p(\beta, Q) = p(\beta)p(Q)$$
$$p(\beta) \sim f_N(\beta|\underline{\beta}, \underline{V_\beta})$$
$$p(Q) \sim f_{IW}(Q|\underline{Q}, \underline{v_Q})$$

For dealing with overfitting I entertain a prior in Minnesota fashion. Prior for β_m (3-month federal funds futures and S&P 500 surprises) is set to 0, other β 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: $\lambda_1 = 0.2$, $\lambda_2 = 0.5$, $\lambda_3 = 1$, $\lambda_4 = 100$. σ_i, σ_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 β is:

$$\beta|y, Q^{-1} \sim N(\overline{\beta}, \overline{V_b})_{I_{s(\beta)}}$$
$$\overline{V_\beta} = (\underline{V_\beta}^{-1} + \sum_{t=1}^{T} X_t' Q^{-1} X_t)^{-1}$$
$$\overline{V_b} = \overline{V_\beta}(\underline{V_\beta}^{-1}\underline{\beta} + \sum_{t=1}^{T} X_t' Q^{-1} y_t)$$

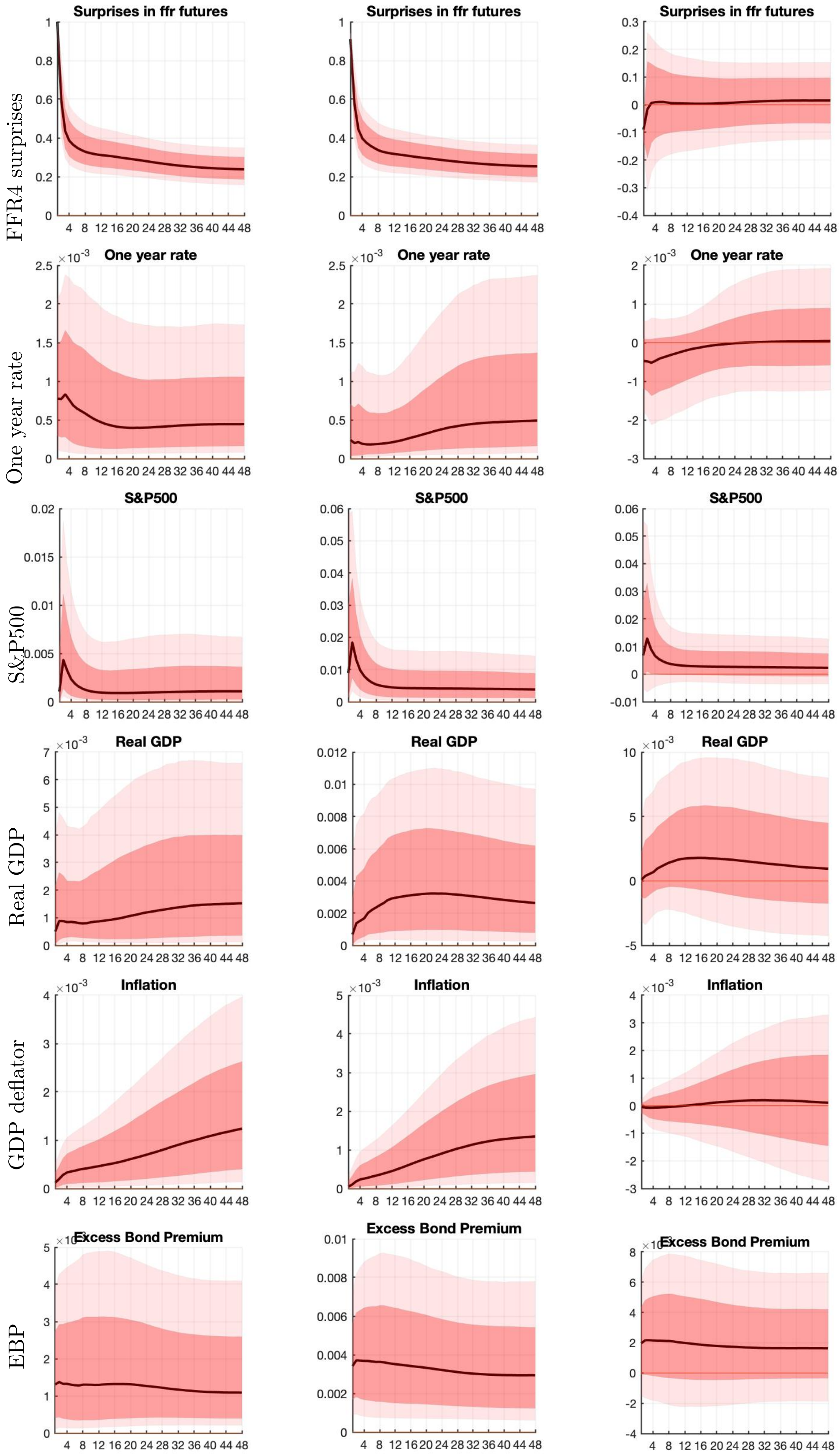
$I_{s(\beta)}$ is an indicator function used to denote that the roots of β lie outside the unit circle.

The conditional posterior of Q is:

$$Q|y, \beta \sim IW(\overline{Q}, \overline{v_Q})$$
$$\overline{v_Q} = \underline{v_Q} + T$$
$$\overline{Q} = \underline{Q} + \sum_{t=1}^{T} (y_t - X_t'\beta)(y_t - X_t'\beta)'$$

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.

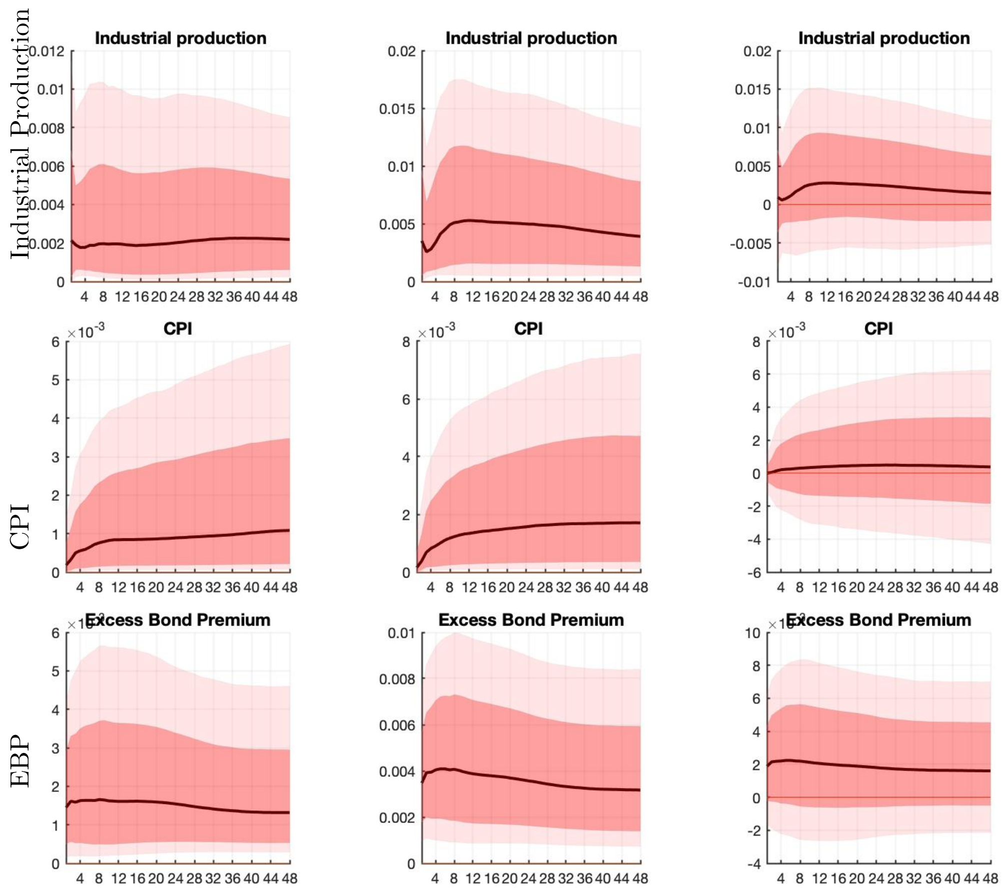
Appendix H. FEVD results



(a) Baseline (b) Policy shock (c) Difference

Figure H.1: Comparison between monetary policy and information shocks

shaded 5%,16%, 84% and 95% percentiles



(a) Baseline (b) Policy shock (c) Difference

Figure H.2: Comparison between monetary policy and information shocks. 3m federal funds futures shaded 5%,16%, 84% and 95% percentiles

Macroeconomic Expectations: News Sentiment Analysis[22]

Nataliia Ostapenko[23]

Abstract

I investigate the role that news sentiment plays in the macroeconomy. Using an approach that combines Doc2Vec embedding and Latent Dirichlet Allocation with lexical-based models I show that the news the media choose to report and the tone of these reports contain important information for household unemployment, interest rates, and inflation expectations. Topic time series derived from the news and the sentiments they express are employed to estimate how the news affects the macroeconomy.

Keywords: expectations, sentiment, news, Latent Dirichlet Allocation (LDA), Doc2Vec

JEL Classification: E52, E31, E00

The views expressed are those of the authors and do not necessarily represent the official views of Eesti Pank or the Eurosystem.

[22] I would like to thank the Bank of Estonia for its hospitality and financial support. I am grateful to Dmitry Kulikov, Vegard H. Larsen and seminar participants at Eesti Pank, 21st IWH-CIREQ-GW Macroeconometric Workshop: Forecasting and Uncertainty for their helpful comments and suggestions, and also to the High Performance Computing Center of the University of Tartu for making the computations feasible. Any errors that remain are my own.

[23] Author's affiliation: University of Tartu, author's email: nataliia.ostapenko@gmail.com. The paper was written while the author was a visiting researcher at the Bank of Estonia.

Appendix I. Introduction

Does the news matter for monetary policy and real economic activity? The idea of the news-driven business cycle suggests that changes in expectations may be an important driver of economic fluctuations. Expectations that future economic conditions will be better, once current fundamentals are controlled for, can be provoked by either news of high TFP in the future, which is known as hard news[24], or positive confidence[25] (**?**). This paper sheds light on the soft news[26] channel of expectations. Since newspapers are the main transmission channel of opinions from professionals to the general public, this type of news might drive changes in household expectations.

An econometrician wanting to identify macroeconomic shocks correctly needs the same information set that decision-makers have, otherwise the identification of the shocks will be subject to omitted variable problems (see **?** for instance). One of these omitted variables might be news about future macroeconomic conditions. Identifying macroeconomic shocks from the data on expectations needs the news about future economic conditions to be controlled for in an empirical model (**?**).

The main data source I use is the business sections of the main US newspapers. News articles are transformed into time series by employing Latent Dirichlet Allocation (LDA) and Doc2Vec embedding with clustering. Afterwards, I adopt a lexical-based approach to assign a sentiment to each article. The lexical approach counts the proportions of positive/negative, constraining and uncertainty words in each article. Combining the topic time series derived from Doc2Vec embedding with clustering and from the LDA with the tone for each news article lets me derive topic time series with sentiments.

These topic time series are employed to identify the types of news that are important for household expectations about unemployment, interest rates, and inflation. I do this

[24]News about economic fundamentals.

[25]Confidence can be viewed as a strong belief that future economic developments will be positive, whereas sentiment is used to describe the views of economic agents about the economic conditions in the future (**?**). Sentiment in a text is a measure of the speaker's tone, attitude, or evaluation of a topic, independent of the topic's own sentiment orientation (Shapiro et al. (2017)).

[26]As opposed to hard news, which means news about objective and directly quantifiable variables such as production and employment, soft news is news with subjective measures of attitudes about current and future economic conditions (Shapiro et al. (2017)).

using a LASSO regression with core macroeconomic indicators from the FRED-MD database (McCracken & Ng (2015)). I reduce the dimensionality of selected news topics with principal component analysis (PCA). The study employs selected topic time series in Structural Vector Autoregressions (SVARs) to overcome a noninvertibility problem (see **?** for further details), and this allows the effects of soft news for monetary policy and real economic activity to be disentangled and the effect of different news sentiments on the macroeconomy to be studied.

The main results are that the Economic topic time series was found to be the most important for household expectations of interest rates, the Housing topic time series mattered most for unemployment expectations, and the Loans topic time series was most relevant for inflation expectations. Moreover, time series of these topics were obtained separately using the two different text transformation approaches of LDA and Doc2Vec with k-means++ clustering. Additionally, the principal component of the time series for these topics was found to have leading properties for indicators of economic activity.

Employing the topic time series mentioned above in conventional VARs with expectation variables showed that a positive soft news shock leads to a long-run increase in real economic activity and consumption, while the effects on inflation and the interest rate are also positive but transitionary. Moreover, the soft news shock accounts for about 20% of the forecast error variance of real economic activity at longer horizons, while the effects of sentiment or expectations shocks were found to be less important. This helps to disentangle the effects of news shocks and sentiment shocks empirically, so it is not necessary to impose ad hoc theoretical identifying assumptions in SVARs.

This study adds to the findings of **?**, Larsen & Thorsrud (2019) and Shapiro et al. (2017) that the transitionary response of inflation to a news shock might be positive. This suggests that news shocks might not be viewed as anticipated exogenous TFP shocks, and that opens the door to alternative news shock channels, such as endogenous growth or anticipated demand shocks with endogenous propagation. This result is in line with the results of **?**, who used expectation variables from surveys in VAR and found that anticipation of economic expansion leads to a fall in unemployment, a rise in inflation, and tighter monetary policy.

On top of that **?** showed that expectations shocks account for a large share of the variance in real economic activity at longer horizons, while the findings from this study suggest that news shocks are more important at longer horizons for economic activity than households' expectations are. Contrary to the findings of **?** this study shows that the news media channel is important for real economic activity and consumption (this is also in line with the findings of Larsen & Thorsrud (2019)).

To identify how soft news affects monetary policy I investigate the effects of the time series for the Fed and Loans topics in an otherwise conventional VAR. It was found that the time series for the Loans topic is more important in the transmission mechanism of monetary policy than that for the Fed topic. That is because households generally do not pay much attention to monetary policy news. At the same time, monetary policy, both conventional and unconventional, can lead to a rise in long-run rates, and so it affects the current decisions of households and firms through this channel. The excess bond premium declines in response to a positive soft news shock, which leads to an expansion in economic activity and tighter monetary policy because of general equilibrium effects.

The results are in line with the findings of Hansen & McMahon (2016), who employed a narrative approach to the FOMC statements to identify what effect forward guidance had. They did not find any significant contribution from forward guidance shocks and such shocks were found to account for a small share of the forecast error variance in real variables. Neither does this study find the Fed sentiments in newspapers to have any strong effect on inflation or economic activity. One possible reason the effects are small is that the paper uses monthly frequency.

Having said that, the findings are in line with those of **?**, who employed high frequency identification and found that monetary policy news about forward guidance did not have any significant effect on household beliefs. Instead they found that news about changes in the target rate has a significant effect on the expectations of households. This is in line with the channel found in the current study, which is that monetary policy news has an effect through changes in long-term rates.

The results of this study add to the results of **?** since the study uses news about monetary policy in its analysis rather than employing survey forecasts as a proxy. **?** found that news about future monetary policy as proxied by survey forecasts has large, immediate and persistent effects on inflation and economic activity. Additionally, previous findings from DSGE models show that monetary policy news shocks are generally more important than unanticipated monetary policy shocks in explaining business cycles (**?**, **?**). The results of this study show that newspapers are not the main channel for these large effects that are observed for anticipated monetary policy.

As an alternative to changing long-run interest rates, the Fed might aim to change the inflation expectations of consumers directly, as was also pointed out by **?**. Changes in household inflation expectations affect the economy through the perceived real interest rate (**?**), but this study did not find empirical support for the direct channel that changes households' inflation expectations being important.

The remainder of the paper proceeds as follows. Section 2 describes different methodologies for sentiment analysis. Section 3 shows the connections between different types of news and household expectations. Section 4 presents ways the selected topic time series can be used to disentangle the effect of soft news on monetary policy and real activity. Section 5 concludes.

Appendix J. Methodology

J.1 Data and sentiment

I use the Nexis Uni database, from which I extract daily business news from The New York Times 1980–2019, The Washington Post 1981-2019, The Los Angeles Times 1985-2019, and The Chicago Tribune 1985-2019. The New York Times is the second-largest newspaper by circulation and the largest circulating metropolitan newspaper, with a weekly circulation of 2.1 million. It is ranked 18th in the world by circulation. The Los Angeles Times is the fourth-largest US newspaper by circulation, The Chicago Tribune is the sixth-largest and The Washington Post is the seventh-largest. The total timespan is 1980:M6–2019:M7.

For comparison, Larsen & Thorsrud (2019) used 25 years of news data, **?** took data from The New York Times and The Washington Post from 1980 to 2000, **?** used 1990–2016, Shapiro et al. (2017) 1980–2018, and **?** 1989–2017.

Following Shapiro et al. (2017) I filtered out the news that does not contain any of the words: said, says, told, stated, wrote, or reported. Imposing this criterion meant the data pull yielded around 416,000 articles.

There are a few mainstream theories about the role news has in the expectation formation mechanism. **?** introduced a noisy-information model, where price-setters get a noisy signal about monetary policy in every period, while **?** model price-setters gaining perfect information about monetary policy with the probability λ in every period, where expectations matter because some price-setters are still setting prices using old decisions and old information. Additionally, some price-setters might learn about monetary policy through a limited-information channel, so it is as if they are observing monetary policy with a random error and have to solve a signal-extraction problem (?). ? pointed out that information costs may make it rational for agents to select methods other than rational expectations, so agents update their previous expectations in each period by weighing costs and benefits.

There are two main channels through which the news affects the economy. The first is reporting on actual economic data and the second is the transmission of opinions from professionals to the general public. Experts express their opinions in the news media and the more important an issue is, the more frequently it is covered by the media. The general public form their expectations from personal opinion and the news, and these expectations influence the current economic decisions of agents. Intensive news reporting improves the accuracy of household expectations because they receive more information (?, ?, ?). The more frequently a story is covered by the news, the more probable it is that households will read it (Larsen & Thorsrud (2019)). However, this effect also depends on the tone of the news (?, ?).

? proposed the following model of the expectation formation mechanism with the media channel (1):

$$M_t(\pi_{t,t+12}) = \alpha_1 N_t(\pi_{t,t+12}) + \alpha_2 M_{t-1}(\pi_{t-1,t+11}) + \alpha_3 P_t(\pi_{t-1}) + \epsilon_t \tag{1}$$

where $\pi_{t,t+12}$ is household expectations, in this case inflation expectations over the next year; $M_t(.)$ is expectations in the current period; $M_{t-1}(.)$ is expectations from the previous period; $N_t(.)$ is news, proxied by ? using a survey of professional forecasters; and $P_t(.)$ is the latest statistics.

? stated that the news affects the perceptions of households through three channels. First it conveys the latest economic data and the opinions of professionals to households; second, it gives them a signal about the economy through the tone and volume of news reporting; and third, the more news about the economy there is, the more likely it is that households will update their expectations about the economy. The authors found evidence that all three of these channels are important for consumer sentiment. Larsen & Thorsrud (2019) tried to capture the effect of the latest economic data in the news, while Shapiro et al. (2017) focused on the opinions of professionals.

The formula quoted above (1) captures the latest economic data in $P_t(.)$ and the opinions of professionals in $N_t(.)$. ? used the survey of professional forecasters directly as a proxy of $N_t(.)$, while I proxy $N_t(.)$ directly by sentiments derived from newspapers. That is in line with **?**, who identified a sentiment shock as orthogonal to surprise and news TFP shocks that maximises the short-run forecast error variance of an expectational variable, which may be a GDP forecast or a consumer confidence index.

Researchers looking at public expectations mainly use time series from the University of Michigan Survey of Consumers as proxies for expectations (**?**, **?**, **?**). The most commonly used surveys are the Survey of Professional Forecasters, the Lundberg Survey, the Michigan Consumers Sentiment Survey, and the Livingston Survey.

As $M_t(.)$ I employ the University of Michigan Survey of Consumers (**?**) and I use expectations of interest rates, unemployment, and inflation over the next 12 months. These correspond to the answers to the survey questions: "No one can say for sure, but what do you think will happen to interest rates for borrowing money during the next 12 months - will they go up, stay the same, or go down?", "How about people out of work during the coming 12 months - do you think there will be more unemployment than now, about the same, or less?", "During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now? and By what percent do you expect prices to go up, on the average, during the next 12 months?".

The importance of expectations of this type has also been pointed out in the earlier literature. **?** estimated a novel measure of the intrinsic value of macroeconomic announcements, which they defined as the ability of the announcement to nowcast GDP growth, inflation, and the federal funds target rate. **?** studied the bond market response to macroeconomic news and grouped announcements into three broad categories of news about prices, news about real output, and news about monetary policy. **?** identified the anticipated and unanticipated components of shocks to technology, demand, and monetary policy using the actual and forecast data for output growth, inflation, and an interest rate.

Larsen & Thorsrud (2019) ran sign adjustment on news topics to separate positive and negative news. As was pointed out by Sims (2003) though, the tone of economic reporting affects sentiment beyond the economic information contained in the reporting itself, as was explored by Shapiro et al. (2017). Therefore, I take in both news frequency and news sentiments.

To assign a sentiment to each news article I employ a combination of two dictionaries by **?** and Loughran & Mcdonald (2016) (LM) with modifications, which are discussed in Appendix A. This approach relates to Shapiro et al. (2017), who found that combining different dictionaries with a negation rule comes closer to human judgement in labelling sentiment. It is worth mentioning that Shapiro et al. (2017) also used a Vader package that was trained on general text labelled by humans from Amazon Mechanical Turk, but the performance of their modified dictionary with Vader was not statistically significantly different from that of the combination of several dictionaries with the negation rule.

Positive sentiment in an article is calculated as the sum of sentences with positive sentiments (2):

$$Pos_i = \sum_{sentence} \frac{\#positivewords_i - \#negativewords_i}{\#totalwords_i} \tag{2}$$

Since news that is more intensively covered is more important, the monthly aggregate positive sentiment for each topic is adjusted by topic frequency within a month. The total monthly positive sentiment for a topic is calculated as the sum of daily positive sentiments minus negative sentiments multiplied by the fraction of all articles covering the topic within a month, or topic frequency (3):

$$Pos_{topic} = \sum_{i \in topic} Pos_i \times \frac{\#topic\_articles}{\#Total\_articles} \tag{3}$$

where $\#topic\_articles$ is the number of articles on one topic within a month, and $\#Total\_articles$ is the total number of articles within that month.

Similarly, I calculated uncertainty and constraining sentiments by employing (2) and (3) for uncertainty and constraining words from Loughran & Mcdonald (2016)[27]. I further use Pos_{topic}, $Uncertain_{topic}$, and $Constraining_{topic}$ as $N_t(.)$ for different news topics in (1). The methodology for extracting the news topics is discussed in the next two subsections.

J.2 Latent Dirichlet Allocation

Following Larsen & Thorsrud (2019) I use the 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 the word order does not matter. Mixing the words in an article and running the LDA will give the same result that not mixing does. Standard text processing steps are employed to extract news topics with the Latent Dirichlet Allocation:

- Words from a stoplist are excluded. This list contains common words that contribute little meaning to the documents, such as prepositions, conjunctions and pronouns.

[27] The full list of words for each sentiment category is available at https://sraf.nd.edu/textual-analysis/resources/

- Words are stemmed, or reduced to their word root form, so economy, economic, economical, economics, economise are all reduced to the root form econom.
- Rare and common words are removed.
- The vocabulary that results consists of 57990 unique words.

The 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 has the following elements (Darling (2011)):

1. For each topic $k = 1, ..., K$, where K is the total number of latent topics:
   - A discrete probability distribution over a fixed vocabulary that represents the k^{th} topic distribution, $\varphi_k \sim Dirichlet(\beta)$[28]
2. For each document $d \in D$, where D is the total number of documents:
   - A document-specific distribution over the available topics (per-document topic proportion), $\theta_d \sim Dirichlet(\alpha)$[29]
   - For each word $w_n \in d$, where N is the total number of words:
     (a) Per-word topic assignment, showing which topic generated the word instance $w_{d,n}$, $z_{d,n} \sim Mult(\theta_d)$[30]
     (b) An observed word, $w_{d,n} \sim Mult(\varphi_k)$

The joint probability for the LDA takes the form (4):

$$p(w_{d,n}, z_{d,n}, \theta_d, \varphi_k|\alpha, \beta) = (\prod_{n=1}^{N} p(z_{d,n}|\theta_d)p(w_{d,n}|z_{d,n}, \varphi_{n,k}))(\prod_{k=1}^{K} p(\varphi_k|\beta))(\prod_{d=1}^{D} p(\theta_d|\alpha))$$

$$= (\prod_{n=1}^{N} Mult(z_{d,n}|\theta_d)Mult(w_{d,n}|z_{d,n}, \varphi_{d,k}))(\prod_{k=1}^{K} Dirichlet(\varphi_k|\beta))(\prod_{d=1}^{D} Dirichlet(\theta_d|\alpha)) \quad (4)$$

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

The latent variables $z_{d,n}$, θ_d, φ_k are unobserved. Inference is done with Collapsed Gibbs Sampling (Griffiths & Steyvers (2004)) with $\alpha = 50$ and $\beta = 0.01$. Since for the inference of both θ_d and φ_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

[28] $Dirichlet(.)$ is the Dirichlet distribution (a conjugate prior for the Multinomial distribution), β is a hyper-parameter

[29] α is a hyper-parameter.

[30] $Mult(.)$ is the Multinomial distribution.

(2004) for the detailed treatment). The outcomes of the algorithm are topic distributions θ_d and word distributions per topic φ_k.

lda_march2.jpg

Figure J.1: LDA topics

The optimal number of topics for LDA was chosen using coherence values. The topics are considered to be coherent if all or most of the words are related, appearing say in the top N words for the topic. Coherence values for different numbers of topics are presented in Figure B1. The coherence values show the optimal number of topics to be 40. All the topics from the LDA model are interpretable and are shown in Figure 1, while Table B1 shows the word distributions for each topic.

J.3 Doc2Vec

Another way to transform articles into a numeric format is the Neural Network (NN) Doc2Vec, which was introduced by **?**. Neural Networks take account of the word order and semantics of words and have no specific text processing requirements. Doc2Vec works with text through stochastic gradient descent and back-propagation. Each paragraph is mapped into a unique vector represented by a column in matrix D, and each word is mapped into a unique vector represented by a column in matrix W. The paragraph vector and word vectors are averaged or concatenated to predict the next word in a given context.

This Neural Network is based on the distributional hypothesis, which is that words that occur in a similar context have a similar meaning. Doc2Vec exploits this hypothesis and transforms words that are similar semantically, as they occur in a similar context, into vectors that are similar geometrically in Euclidean space. Doc2Vec transforms articles into vector representations, and the representations for conceptually similar articles are close to each other by cosine similarity. Doc2Vec does not rely on the bag-of-words assumption, so word order matters for it.

The objective of Doc2Vec is to maximise (5)

$$\frac{1}{T}\sum_{t=k}^{T-k}\log p(w_t|w_{t-k},...,w_{t+k}) \tag{5}$$

$$p(w_t|w_{t-k},...,w_{t+k}) = \frac{e^{y_{w_t}}}{\sum_i e^{y_i}}$$

$$y = b + Uh(w_{t-k},...,w_{t+k}|W,D)$$

where $w_{t-k},...,w_{t+k}$ is a sequence of words, $p(.)$ is a probability (softmax), y_{w_t} is the unnormalised log-probability for each output word, b and U are the softmax parameters, and $h(.)$ is constructed by concatenation or averaging of the word vectors extracted from W (for further details please consult **?**).

I normalised the vectors to have unit lengths. In this case, minimising Euclidean distance is the same as maximising cosine similarity (see Appendix C for details).

For topic clustering I employ k-means++ (**?**). The task is to take a set of n points in Euclidean space and an integer k and find a partition of these points into k subsets, each with a representative, also known as a centre. The k-means method seeks to minimise the average squared distance between points in the same cluster.

k-means++ chooses initial clusters in a different way to k-means clustering and afterwards proceeds to standard k-means clustering. Given a set of observations $(x_1, x_2, ..., x_n)$, $x \in X$, k-means clustering aims to partition the n observations into k sets $S = \{S_1, S_2, ..., S_k\}$ to minimise the within-cluster sum of squares (3):

$$L = \underset{S}{argmin} \sum_{i=1}^{k} \sum_{x \in S_i} ||x - \mu_i||_2^2 \tag{3}$$

where μ_i is the centre of a cluster. The task of the algorithm is to find the set of centres (μ_i) that minimises the objective function (3).

k-means clustering iterates over two steps until convergence:

1. Assign each input point to its nearest centre.
2. Recompute the centres given the point assignment.

While k-means uses a random initialisation of cluster centres as a local search, k-means++ selects only the first centre uniformly at random from the data. The probability of each subsequent centre being selected is proportional to its contribution to the overall error given the previous selection. The probability of a point being chosen to be the i^{th} centre depends on the realisation of the previous centres. An implementation of k-means++ initialisation will make k passes over the data to produce the initial centres (**?**).

k-means++ proceeds as follows (**?**):

1. Take one centre μ_0, chosen uniformly at random from X.
2. Take a new centre μ_i, choosing $x \in X$ with probability $p = \frac{D(x)^2}{\sum_{x \in X} D(x)^2}$[31]
3. Repeat Step 2 until the algorithm has taken k centres altogether.
4. Proceed as with the standard k-means algorithm (steps 1-2 from the standard k-means algorithm until convergence).

I calculate the most common words from the news headlines for each cluster. The results for 40 clusters are presented in Figure 2. Table D1 shows the detailed topics' descriptions from Doc2Vec with k-means++ 40 clusters.

[31] $D(x)$ is the shortest distance from a data point to the closest centre that has already been chosen.

Doc2vec_march2.jpg

Figure J.2: Topics according to Doc2Vec

Appendix K. Soft news

K.1 Topic time series with sentiments

I employ two different methods for assigning topics for the LDA results since the model gives topic distributions as an output. The first is to assign a dominant topic for each article[32] and the second is to assign topic distributions for each article[33]. k-means++ clustering assigns one topic for each article, so I have three topic time series models in total.

Figure E1 shows the cross-correlations between the topic time series with the sentiments from Doc2Vec with k-means++. The topics that correlate most are Financial markets and Dow, Investment and Financial markets. The figure also includes household expectations. The highest correlations with expectations of interest rates were found for unemployment expectations, and the topics Reports, Economic, Profits, Money, Company stocks, Retailers and Jobs; for consumer unemployment expectations the correlations were with interest rates expectations, and the topics Vehicles, Economic, Housing, Financial market and Reports; and those for inflation expectations were with the time series for the topics Housing and Loans.

Figure E2 presents the correlations between the topic time series for the sentiments from Doc2Vec with clustering and the LDA model that assigns the topic frequency for each article. Although methods employed are different, many topics are correlated between different models, such as the time series for the topics Housing, Fraud and Law, Dow and Stocks, Jobs, Profits, Currency, Company stocks and Deals, Retailers, Airlines, Economic and Reports, Loans, Oil/gas, Vehicles and Cars, Investing, President, Technology, and Computers.

Figure E3 discusses the correlations between the topic time series for the sentiments from Doc2Vec with clustering and the LDA model that assigns a dominant topic for each article. There are meaningful correlations between the time series for the topics Dow and Stocks, Housing, Profits, Jobs, Currency, Retailers, Energy, Airlines, Financial market and Stocks, Banking and Credit, Economic and Reports, Oil/gas, Services, Real estate and Cities, Loans, Vehicles and Cars, Health, Financial news and Securities, Investing, Fed, Technology

[32]Which is similar to clustering as one article is connected with the one topic that has the highest proportion among the 40 topics.

[33]One article is associated with 40 topic proportions.

and Computers, and President.

un3.png

(a) Unemployment positive sentiment in red and the topic Economic (positive/negative) in blue

un1.png

(b) Unemployment positive sentiment in red and the topic Jobs (uncertain/certain) in blue

ffr.png

(c) FFR positive sentiment in red and the topic Economic (positive/negative) in blue

Figure K.1: Comparison of the LDA topic time series with sentiment and simple sentiment frequency models. All series are standardised.

Figure E4 shows the correlations between the time series of topics for the sentiments derived from differently labelled LDA models. Almost all the time series that represent the same topic are highly correlated. Appendix E presents the dynamics of the topic time series with sentiments from different models. Even though the assumptions of the underlying models are different, the topic time series derived from the different models have quite similar dynamics and similar labels.

For comparison, I extracted sentences with keywords related to unemployment, inflation and the federal funds rate (FFR)[34], and calculated a sentiment for each keyword using the methodology described in Appendix A. After each keyword was detected, the sentiment of positive or negative was assigned to a sequence of the five words preceding it and the five words following it. The sentiment related to a specific keyword, if any, should appear in this window. I add up these sentiments for each keyword for each month. The comparisons between the sentiment models from keywords and the topic models with sentiments are shown in Figure 3.

Surprisingly, the correlation between the frequency of positive sentiments for unemployment and positive sentiments for the time series of the topic Economic from the LDA model using topic frequencies as labels is 0.57[35]. That is because the tone of the news is more negative during recessions and so the negative sentiments are more common.

The uncertainty sentiment for the time series of the topic Job is negatively correlated with the frequency of positive sentiments for the keyword unemployment (part (b) of Figure 3). During recessions there is greater uncertainty about unemployment and negative news articles about unemployment appear more frequently.

Similarly, it is seen that positive news about the federal funds rate is more common during bad times (part (c) of Figure 3). That is completely in line with the objectives of the Federal Reserve, or the Fed, as it revises rates downward to stimulate the economy during recessions. Moreover, the Fed signals more about monetary policy during recessions.

[34] The keywords for the federal funds rate are discount, rate and federal. A sentence should contain at least two of these keywords.

[35] All correlations in this section are presented from 1984 because of the high level of volatility in economic variables during the Volcker disinflation.

fred.png

(a) The first factor in red and the topic Economic in blue

fred2.png

(b) The first factor in red and the topic Loans in blue

fred3.png

(c) The first factor in red and the topic Housing in blue

Figure K.2: LDA topic time series with positive sentiments and the first factor from the FRED-MD (McCracken & Ng (2015)). All series are standardised.

Figure 4 shows the relation between the topic models and the first factor from the FRED-MD database[36] (McCracken & Ng (2015)). The factor has strong correlations with the time series with positive sentiments of 0.65 with the topic Economic, 0.55 with the topic Loans, and 0.62 with the topic Housing. Since the first factor from the FRED-MD is usually used as an indicator of business cycles in the economy, these correlations between the factor and different topic time series cannot occur just by chance.

Parts (b) and (c) of Figure 4 discuss the connection between the time series with positive sentiments for the topics Housing and Loans and the first FRED-MD factor. These topic time series are correlated with business cycles since mortgage and loan rates tend to fall during recessions, as do house prices. Additionally, these topics might be covered more frequently during recessions, or might be more negatively framed by experts. This then means that there is asymmetric news coverage in bad times and good times. People might also respond differently to bad news and to good news (**?**). This is also confirmed by the results of **?**, who found that the absolute forecast error of people who hear news about higher inflation is much higher than that of people hearing news about lower inflation.

K.2 Expectations and soft news

The next question is whether the topic time series with positive sentiments that are derived contain additional information for household expectations. **?** employed the least absolute shrinkage and selection operator (LASSO) with the news topics and the FRED-MD database, which contains major macroeconomic indicators for the US economy (McCracken & Ng (2015)). In line with that, I employ LASSO together with 125 stationary monthly variables from FRED-MD. The FRED-MD variables should capture the hard news channel since hard news should contain information about economic fundamentals.

This study aims to capture the effect of the news as a transmission channel since professionals usually follow hard economic indicators and express their opinion about current and future economic developments to the public through the media. I proceed by checking which of the news topic time series with sentiments are important for household expectations. For the LASSO estimation I employ (1), which is presented again here for convenience:

$$M_t(\pi_{t,t+12}) = \alpha_1 N_t(\pi_{t,t+12}) + \alpha_2 M_{t-1}(\pi_{t-1,t+11}) + \alpha_3 P_t(\pi_{t-1}) + \epsilon_t$$

[36]This factor extracted from the FRED-MD data set is obtained using the Principal Components Analysis (PCA).

where $\pi_{t,t+12}$ is inflation expectations over the next year, $M_t(.)$ is expectations in the current period, $M_{t-1}(.)$ is expectations from the previous period, $N_t(.)$ is the news proxied by **?** using the survey of professional forecasters, and $P_t(.)$ is the latest statistics.

As $M_t(.)$ I use consumer expectations from the **?**, which gives expectations for interest rates, unemployment and inflation. In some specifications I add consumer expectations from the previous period to capture $M_{t-1}(.)$. The FRED-MD macroeconomic indicators from the previous period should capture the latest statistics available to agents in $P_t(.)$. $N_t(.)$ are topic time series with sentiments from the previous month to avoid the problem of simultaneity. If I used N_t from the same month, it would contain information not available to the agents during the surveys[37].

The regularisation parameter for LASSO is chosen using a five-fold cross-validation. All the non-stationary series were transformed into a stationary form by taking first differences. Additionally, all the variables were standardised. This is done since LASSO tends to select one variable from among highly correlated ones and from many covariates it selects those that have a large effect. Standardisation therefore makes LASSO invariant to scale.

Appendix F presents the LASSO results for expectations for interest rates, unemployment and inflation with the FRED-MD variables and the topic time series with sentiments from different models. The columns differ in how the sentiments were assigned to different topic time series. The first column shows the results for positive sentiments for different topics, the second column shows those for uncertain sentiments for topics, and the third column shows them for constraining sentiments for topics. The first three columns do not include $M_{t-1}(.)$, the expectations from the previous period. All the other columns include expectations from the previous period as a control. The seventh column presents the LASSO results for interaction between positive and uncertain sentiments and the last column illustrates the results for the interaction of positive and constraining sentiments for each topic. All the columns include the FRED-MD variables, though outputs are omitted for the sake of brevity.

1. The LASSO results for interest rate expectations. The Fed topic time series is found to be connected to household interest rate expectations (Table F1 columns 4,7,8, Table F4

[37]For example, some consumers were surveyed between 2 and 10 January. The news time series for a month is the sum of sentiments during the whole month, so it contains more information than was available to consumers on 10 January. A survey might also start at the end of the previous month and finish at the end of the current month.

column 7 and Table F7 columns 4,7,8). This topic might be important for the interest rate expectations since it covers news about monetary policy and the Federal Reserve, which might give more information about monetary policy during recessions or during periods with a zero lower bound. The Federal Reserve may also signal the future path of monetary policy to the public, for instance when it acts under commitment. It might equally use the media to transmit a signal for forward guidance, but if households were tracking and reacting to the federal funds rate hour by hour, it would not matter whether newspapers cover the topic in depth or not at all (Sims (2003)).

The results also show how important the Economic topic time series is for interest rate expectations (Table F1 column 1, Table F4 all columns and Table F7 all columns). The time series for the topic Economic might be connected to household interest rate expectations since it captures general information about the economy. Negative news about the economy is more common in recessions, while the media cover more news about other topics during expansions. Interest rate expectations are also generally higher during recessions.

2. The LASSO results for unemployment expectations. The President topic time series has non-zero coefficients in LASSO regressions for household expectations about unemployment (Table F2 columns 3,4,5,6, Table F5 column 4, and Table F8 columns 4,6). This topic might reflect general expectations about future economic conditions. The topic President might occur more frequently during bad times, and less often during good times. The tone of the President's statements might also be correlated with business cycles.

The Jobs topic time series with uncertain sentiment was found to be positively connected with unemployment expectations (Table F2 column 5 and Table F5, columns 5,7,8). This might suggest that households pay more attention to the topic when there is economic uncertainty.

The Housing topic time series is the most important for unemployment expectations (Table F5 all columns and Table F2 columns 1,3,4). The Housing topic time series with positive sentiment is negatively connected with unemployment expectations. This argument is supported by earlier studies that have found a negative correlation between regional labour and housing markets during the 2007–2009 recession. The previous studies also found a relationship between housing prices and unemployment, which might arise because the housing supply is inelastic (?). An alternative explanation is that this time series is related to business cycles in the US economy. Moreover, over two-thirds of households in the US own houses and invest the majority of their portfolio in real estate (?). This means households are likely to pay a lot of attention to house prices. A similar argument might be given for the importance of the Loans topic time series from the LDA model using dominant topic labels (Table F8 columns 1,3,4,6,8).

3. The LASSO results for inflation expectations. Table F3, Table F6, and Table F9 show the results for inflation expectations and the topic time series from different models. Among all the models the time series for the topic Oil/gas was found to add additional information alongside economic fundamentals and the past inflation expectations of households. This result is not surprising since oil prices make up the largest part of gasoline prices. Americans generally drive cars and so must pay attention to gasoline prices. Households pay less attention to newspapers and more attention to gasoline and retail prices (**?**).

The Loans topic time series from the LDA models was found to be the most important for inflation expectations (Table F6 all columns and Table F9 columns 1,3,4,6,8). Long-term interest rates rise during bad times, and the Federal Reserve can change long-term rates through its communication and the federal funds target channel. These channels change the inflation expectations of professionals and policy-makers, and this in turn leads to changes in the long-run rates. Households might follow information about the long-run rates more closely since this information is important for their current economic decisions.

To study how the news affects business cycles, Larsen & Thorsrud (2019) constructed a news index from topic time series as a weighted average of news topics with the highest predictive score. Similarly, **?** weighted selected topics by partial R^2 from OLS results based on topics selected by LASSO. The authors developed a news index as a linear combination of news topics weighted by their relative importance for expectations. **?** employed principal component analysis (PCA) to reduce the dimensionality of the five least connected topics, which are more likely to be exogenous.

In line with this, I employ PCA to find how the news affects the macroeconomy. PCA is a method for extracting features and reducing dimensionality, as each component captures the direction of the maximal variance of the data and each component is orthogonal to every other component. These PCAs might be used in the same manner as factors from the FRED-MD database to augment the standard Vector Autoregressions with additional information variables (see **?** for example).

Since households are unlikely to follow the latest macroeconomic statistics but will follow the news, the first principal component from the news might help to identify the effects of anticipated macroeconomic shocks. Figure 5 shows the first principal component of positive sentiments for the time series of the topics Economic, Housing, Loans, and Oil/gas from the LDA model using topic distribution labels. As was found earlier in this study, the Economic topic time series is important for consumer expectations of the interest rate (Table F4), the Housing topic time series is important for consumer expectations of unemployment (Table F5), and the Loans topic time series is important for consumer expectations of inflation (Table F6). The Oil/gas topic time series is also important for households' expectations of inflation, and as was pointed out by **?** moreover, changes in gasoline prices might lead to changes in consumer expectations of inflation. Indeed oil price shocks were one of the major drivers of US inflation from 1973 (**?**).

On top of that, the previous findings of **?** can validate the choice of variables. **?** used high-frequency identification and found that surprises in jobs reports, GDP reports and housing starts releases affect consumer confidence.

The principal component extracted moves in tandem with the first factor from the FRED-MD, which suggests that it does not capture just noise. Moreover, it has leading properties with respect to the first factor from the FRED-MD. Figure F.1 presents the factor loadings and Figure F.2 compares the first principal component from the topic time series with the first factor from the FRED-MD database.

factor_pca.png

Figure K.3: The first factor from the FRED-MD in blue and the principal component from news topic time series in red. All series are standardised.

The topics are Economic, Housing, Loans, Oil/gas. Positive sentiment.

Shaded areas - NBER based Recession Indicators for the United States

Appendix L. The role of news sentiments

L.1 Soft news and economic activity

To identify how soft news affects the real economy I use the same data as Shapiro et al. (2017) for the period 1984:M1–2019:M7. These data are the logarithm of industrial production, the logarithm of real personal consumption expenditures, the logarithm of the PCE price index, and the federal funds rate[38]. All the data are obtained from the *Federal Reserve Economic Data* (2019). Additionally I employ the consumer sentiment index from the **?**. I use the first factor from McCracken & Ng (2015) as a measure of hard news from current economic indicators to disentangle the effect of the soft news channel. As soft news I use the first principal component from news defined in the previous section. The principal components were standardised. I use twelve lags because the data are at monthly frequency. Details of the estimation and priors can be found in Appendix G.

Figure 6 presents the results with two alternative ordering schemes. The first has hard news $(t-1)$, soft news $(t-1)$, output, consumption, inflation and the real rate, and the second ordering scheme has the soft news variable ordered last. These alternative ordering schemes represent different assumptions about a structural news shock. In the first scheme the structural shock affects output, consumption and the real rate on impact, whereas in the second scheme it affects only soft news on impact.

Considering that soft news should capture only news sentiments, which are the opinions of professionals rather than reports on current macroeconomic indicators, I control for the current macroeconomic indicators in VAR. To separate the effects of these two components, hard news, which is the first factor from the FRED-MD, is ordered before soft news.

The underlying mechanism of news shocks should be that when households and firms become more optimistic about future economic prospects, they start to spend more, so consumption and investment increase, driving up aggregate activity. If the optimism of the agents turns out to be justified, the economy converges to a higher long-run path; if it does not, the economy returns to its original trend because of general equilibrium forces (**?**, **?**, **?**,

[38]Using the federal fund rate minus expected inflation from the **?** gives similar empirical results, except that a soft news shock accounts for a larger share of the forecast error variance of the real rate (see Figure H.3).

Larsen & Thorsrud (2019) [39]).

The findings (see Figure 6) support the expected effect of news shocks as output and consumption start to increase sluggishly in response to a one standard deviation soft news shock and converge to a new long-run equilibrium. The real interest rate starts to rise due to general equilibrium effects in response to increasing output and consumption.

10000.jpg

(a) Soft news shock, ordered second

last.jpg

(b) Soft news shock, ordered last

Figure L.1: Impulse responses to soft news shocks

median and 16th and 84th percentiles

Contrary to the previous findings by **?**, Larsen & Thorsrud (2019), and Shapiro et al.

[39] According to these authors a news shock is a shock that increases expected future productivity without affecting current productivity.

(2017) though, there is a transitory rise in inflation in response to a soft news shock. This effect can be explained by demand shocks with an endogenous information structure (**?**) or an endogenous growth mechanism (**?**). At the same time, the findings of this study are in line with **?**, who found that positive expectations typically lead to a significant rise in economic activity, inflation, and the interest rate.

Since soft news is ordered after hard news, it is orthogonal to the current macroeconomic indicators and should contain only subjective information about future productivity, maybe with some noise. The forecast error variance decomposition (FEVD) (Appendix H) shows a soft news shock accounts for about 15% of the variance in output at longer horizons and up to 10% of the variance in hard news. **?** found that the news media are not an important channel for consumer confidence, but rather that consumers aggregate information from different sources. This claim about how the actual news affects future economic developments can be tested using the data from this paper. Figure 7 presents the results with the Consumer Sentiment Index from the **?** as an additional measure of consumer confidence. It is ordered before soft news since the topics were selected for their importance for consumer expectations at time $t + 1$, so the news should affect consumer expectations in the next period. In this case a sentiment shock[40] is a structural shock that affects consumer expectations and news on impact, whereas a news shock only affects soft news on impact. The expectation variables were standardised.

It should be noted that I employ the FRED-MD factor, ordered before soft news, to purge the soft news from any reports on hard macroeconomic indicators, so that it captures only sentiments. For contrast, Figure 7 also presents the impulse responses of variables to the first structural shock, which influences all the variables included on impact. In this case the Fed reacts more aggressively in the long-run, and that reaction restrains the positive response of inflation for a few months after the shock. There are also no prolonged responses to this shock from output or consumption.

[40] A sentiment shock represents shifts in expectations about business cycles without changes in the fundamentals of the economy (**?**).

soft.jpg

(a) Soft news shock, ordered last

hard_news.jpg

(b) First shock

sentiment.jpg

(c) Sentiment shock, ordered second last

Figure L.2: Impulse responses to soft news and sentiment shocks
median and 16th and 84th percentiles

The impulse responses to a soft news shock are approximately the same as in Figure 6 (b) though the transitionary response of prices is somewhat larger in magnitude. The forecast error variance decomposition shows that the contributions of a soft news shock at longer horizons are also similar to the previous findings (Figure H.2 (a)).

At the same time, the impulse responses to a sentiment shock are similar to **?**, except that the response of output is more uncertain. However, according to the FEVDs, the sentiment shock does not account for a large share of the variance for output and consumption at long horizons (Figure H.2), as it only accounts for up to 10% of the variance of soft news at high frequencies. In this the findings of this study do not support those of **?**, who argued that newspapers are unlikely to be an informational channel for household confidence.

The soft news shock identified cannot be an animal spirit shock[41] since an increase in the real rate in response to a positive but false signal should dampen it. The impulse responses show, however, a persistent effect on output and consumption (Figure 6 and Figure 7) and the FEVDs confirm its importance at longer horizons (Appendix H).

The argument of **?** relies on the assumption of exogenous growth in technology. Their findings do not indicate any causality of news, but rather the perfect foresight of future technological development. Similarly, Larsen & Thorsrud (2019) considered the process of forming expectations as a signal extraction problem, where one part of the signal is the true state of future TFP. The results of this study might uncover a possible endogenous growth mechanism, as demand shocks might cause a short-run increase in real activity, which might ultimately lead to a rise in TFP through an endogenous mechanism of learning by doing and similar.

Employing consumer expectations in similar settings, **?** also found positive co-movement between real economic activity, the real rate and inflation in response to a positive expectation shock. The authors interpreted it in a framework of an expected positive demand shock with search and matching frictions. Similarly, **?** found positive correlations between sentiment, future economic development, and consumption by households.

L.2 Heterogeneity of soft news shocks

Next is the issue of whether different types of news have similar effects for the macroeconomy. It is unlikely. It is possible though to disentangle the effects of different types of news from

[41]**?** refer to animal spirits as false news.

different types of household expectation.

Figure 8 presents impulse responses to each type of soft news shock [42]. The impulse responses to different types of consumer sentiment shock[43] are shown in Figure I.1. Different types of soft news are ordered last, and consumer expectations are ordered second last in each specification. I define a soft news shock as a structural shock that affects soft news on impact and a sentiment shock as a shock that affects consumer expectations and soft news on impact. The news and expectation variables were standardised.

In contrast to the results of **?**, the response of inflation to a positive sentiment shock is negative (Figure I.1 (a)). Neither do the contributions of this shock to the forecast error variances support the previous findings of **?** (Figure I.2 (b)).

The impulse responses to soft news shocks about housing and loans are quantitatively similar, but the response of inflation to a news shock about loans has a higher magnitude for a few months after the shock. The impulse responses to an economic news shock are somewhat tighter but are in the same direction.

Figure I.2, Figure I.3, and Figure I.4 show the contributions of each shock to the forecast error variances of the variables. The housing news shock accounts for about 20% of the variance in output at the horizon of thirty months, while a loans news shock accounts for 37% and an economic news shock for about 7%. Housing and loans news shocks each contribute about 5% of the variance in consumption at the horizon of thirty months.

These different types of news should capture the news sentiment effects in the VARs since none of the consumer sentiment shocks account for variances in real variables at longer horizons, except interest rate sentiment. The interest rate sentiment shock accounts for about 17% of the interest rate error variance. In this case the sentiment shock identified might be an animal spirit shock, since **?** found that this shock has very little effect on the real variables within the exception of the real interest rate.

[42]From the point of view of identifying restrictions all these shocks are soft news shocks as each of them affects only soft news contemporaneously and reacts to all other variables. The label "some news shock" is given here for convenience.

[43]The previous note also applies here.

housing.jpg

(a) Housing news shock, ordered last

soft_loans.jpg

(b) Loans news shock, ordered last

soft_ec.jpg

(c) Economic news shock, ordered last

Figure L.3: Impulse responses to soft news shocks

median and 16th and 84th percentiles

? studied how unemployment expectations affect business cycles. They found positive co-movement in economic activity, inflation and the interest rate in response to expectation shocks, with the expectations variable ordered first in the VAR. The authors also found these shocks made a significant contribution to the forecast error variances of unemployment and inflation at long horizons. If the news is omitted from the VARs though, expectations might falsely account for a large share of the FEVDs of the real variables.

In line with the last argument, Figure I.5 compares the impulse responses to an oil/gas news shock and an inflation sentiment shock. The Oil/gas topic time series was not found to be robust for inflation expectations in all the LASSO specifications and so its connection to expectations might be weaker. The impulse responses to an oil/gas news shock are seen to be different from the IRFs in Figure 8. Additionally, a sentiment shock accounts for about 15% of the variance in output and 7% of the variance in consumption at longer horizons (Figure I.6), while the oil/gas news shock does not account for a significant part of the forecast error variances of the real variables. This finding contradicts the previous findings (Figure I.3).

L.3 Soft news and monetary policy

Lastly, I employ the news time series to study how news sentiments affect monetary policy to disentangle the role of the media. The standard framework for studying monetary policy is to employ recursive identification in three variable Structural Vector Autoregressions with variables measuring economic activity, inflation and a monetary policy indicator. I use additional news variables that were found to be important for consumer expectations of the interest rate.

Recent studies have pointed out that Fed announcements might contain information on the general economic outlook (see Jarocinski & Karadi (2020) and **?** among others). The news about monetary policy does the same. To disentangle the effect of sentiments about the general economic outlook from that of sentiments about monetary policy, I also control for the Economics topic time series, since it was found to be important for consumer expectations of the interest rate, and it was also highly correlated with the first factor from the FRED-MD database that captures general economic conditions.

I use the logarithm of industrial production (IPB50001N) (alternatively the index of real economic activity (CFNAI)[44]) as a measure of economic activity, the logarithm of the

[44]CFNAI aggregates information from a panel of 85 macroeconomic time series encompassing four types, or groups,

consumer price index (CPIAUCNS), the one-year constant-maturity Treasury yield as a monetary policy indicator (GS1)[45], and the excess bond premium as an indicator of financial conditions (EBP[46]). All the data except for the EBP are obtained from the *Federal Reserve Economic Data* (2019).

Variables are ordered in the order given above, followed by the Economic topic time series and the Fed topic time series[47]. The timing also supports the choice of ordering as the news data are aggregated over the current month and so it is plausible that they might react to changes in economic activity or to Fed actions within the current month; moreover, the news over the current month does not affect employment and inflation contemporaneously. This assumption follows from the notion that prices and employment are slow to adjust. I employ twelve lags since the data are monthly. The timespan is 1984:M1–2016:M8. The details about estimation and priors can be found in Appendix G. The last shock is labelled as a Fed news shock[48] and identified as a shock that affects only the Fed topic time series contemporaneously.

Figure 9[49] shows that the one-year rate starts to rise in response to a positive Fed news shock one month after the shock. Real economic activity gradually increases in a few months after the shock, then the speed of increase accelerates ten months after the shock and it finally reaches a new long-run equilibrium about fifteen months after the shock. The shock has a transitory negative effect on inflation, which declines for a few months after the shock.

of indicators: production and income; employment, unemployment, and hours; personal consumption and housing; and sales, orders, and inventories. More information can be found at https://www.chicagofed.org

[45]Since it incorporates the impact of forward guidance and remains a valid measure of the monetary policy stance even when the federal funds rate is constrained by the zero lower bound (Jarocinski & Karadi (2020)).

[46]This variable aggregates high-quality forward-looking information about the economy. The EBP is a component of corporate bond credit spreads that is not directly attributable to expected default risk and provides an effective measure of investor sentiment or risk appetite in the corporate bond market (**?**). The excess bond premium represents credit supply conditions.

[47]In this section I use positive sentiments of the topic time series.

[48]From the point of view of theoretical restrictions this shock is still a soft news shock since it affects only soft news contemporaneously and reacts to all other variables. The label "Fed news shock" is given here for convenience.

[49]Appendix J discusses the results from additional SVAR specifications. Since 2008 the Federal Reserve has relied on unconventional monetary policy measures because the federal funds rate hit the zero lower bound. Figure J.3 shows the result for the sub-period of forward guidance.

The news shock also leads to a decline in the excess bond premium.

1ip.jpg

13.png

(a) Impulse responses to a Fed news shock with uncertainty

(b) SVAR with CFNAI

Figure L.4: Impulse responses to a Fed news shock (from Doc2Vec topic time series)
median and 16th and 84th percentiles

The responses are somewhat similar to those for the central bank information shock of Jarocinski & Karadi (2020), which is similar to an anticipated demand shock that the central bank partly offsets. A decline in the excess bond premium after a Fed news shock indicates an expansion in the supply of credit. This might suggest that the Fed topic time series indicates an endogenous response by the Fed to an anticipated demand shock.

Figure J.4 shows the contributions of the fifth and sixth shocks, labelled as economic news and Fed news shocks, to the forecast error variances of the macroeconomic variables. It is apparent that the economic news shock contributes substantially to the Fed topic time series from the beginning, while the Fed news shock accounts for only 2% of economic activity at longer horizons.

An explanation might be in how much attention households pay to news about monetary policy. **?** documented that households and firms do not generally follow even large policy change announcements, despite widespread news coverage. Only professionals pay attention to monetary policy announcements, while households mainly rely on their prior beliefs.

Conventional and unconventional monetary policies primarily influence the economy through their effects on long-term interest rates (**?**) and long-term interest rates are most important for households' spending decisions. In line with that, **?** also noticed that professionals closely follow the Federal Reserve announcements, and they might change contemporaneous long-term interest rates through financial markets. The finding of the connection between the loans topic time series and inflation expectations might confirm that this is a possible transmission mechanism for monetary policy.

To test this, I employ the Loans topic time series instead of the Fed topic time series in the settings described above. The impulse responses are similar to those presented in Figure 10, but are of higher magnitude. The only difference lies in the response of inflation, which increases after the shock.

10000_fed.jpg

Figure L.5: Impulse responses to a Loans news shock
median and 16th and 84th percentiles

Figure J.5 presents the contribution of a loans news shock to the forecast error variances of the variables. The figure reveals that the shock explains about 20% of the variance of industrial production at the horizon of thirty months and about 5% of the variance of the excess bond premium.

To investigate further how a soft news shock affects monetary policy I added inflation expectations from the **?** and ordered it before the Loans topic time series. The impulse responses to the soft news shock are very similar to those in Figure 9, and the FEVDs of both shocks are presented in Figure J.6. The soft news shock is quite exogenous in this setting and explains about the same share or more of the variance of industrial production in thirty months. From the tenth month this shock explains around 10% of the variance in inflation expectations. The sixth shock, which contemporaneously impacts inflation expectations and the news about loans, explains up to 7% of the variance of industrial production at all horizons.

Appendix M. Conclusions

The study combines the techniques of Doc2Vec with clustering, LDA, and lexical methods to transform the data from newspapers into topic time series with sentiments. The findings show that the Economic topic time series is connected to household expectations for the interest rate, the Loans topic time series is connected to inflation expectations, and the Housing topic time series is connected to unemployment expectations. By combining these topic time series with the Oil/gas topic and reducing the dimensionality, the study derives an indicator of news sentiments about business cycles. This indicator has leading properties for the business cycle indicator based on official statistics.

The first principal component from the positive sentiments of the Loans, Housing, Economic, and Oil/gas topic time series is employed in Structural Vector Autoregressions to identify the role of soft news, which means the news that presents the subjective opinions of experts about the future development of the economy. The study finds that a soft news shock accounts for about 20% of the forecast error variance of output at long horizons. Decomposing the principal component by soft news shocks to separate topics accounts for about 7-27% of the forecast error variance of output and about 5% of the variance in consumption at long horizons in different models. Moreover, the inclusion of news variables leads consumer sentiment and expectation shocks to play a smaller role in SVARs. The effect of a positive soft news shock is in line with an expected positive demand shock with an endogenous propagation mechanism.

On top of that, the study finds empirical support that the transmission mechanism for monetary policy lies in the effect on the long-run interest rates. Households do not pay much attention to the news about monetary policy, whereas the topic time series about loans is important for their inflation expectations.

References

Arthur, D. & Vassilvitskii, S. (2007), 'k-means++: The Advantages of Careful Seeding', *SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms* pp. 1027–1035.

Bahmani, B., Moseley, B., Vattani, A., Kumar, R. & Vassilvitskii, S. (2012), 'Scalable K-Means++', *Proceedings of the VLDB Endowment* **5**(7), 622–633.

Banbura, M., Giannone, D. & Reichlin, L. (2010), 'Large Bayesian vector auto regressions', *Journal of Applied Econometrics* **25**(1), 71–92.

Barsky, R. B. & Sims, E. R. (2012), 'Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence', *The American Economic Review* **102**(4), 1343–1377.

Beaudry, P. & Portier, F. (2006), 'Stock Prices, News, and Economic Fluctuations', *The American Economic Review* **96**(4), 1293–1307.

Beaudry, P. & Portier, F. (2014), 'News-driven business cycles: Insights and challenges', *Journal of Econometric Literature* **52**(4), 993–1074.

Beckers, B. (2020), 'Credit Spreads, Monetary Policy and the Price Puzzle', *Research Discussion Paper Economic Research Department Reserve Bank of Australia* (2020-01).

Beechey, M. J. & Wright, J. H. (2009), 'The High-Frequency Impact of News on Long-Term Yields and Forward Rates: Is It Real?', *Journal of Monetary Economics* **56**(4), 535–544.

Benhabib, J. & Spiegel, M. M. (2019), 'Sentiments and Economic Activity: Evidence from US States', *The Economic Journal* **129**(618), 715–733.

Bernanke, B. S., Boivin, J. & Eliasz, P. (2005), 'Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach', *The Quarterly Journal of Economics* **120**(1), 387–422.

Bernanke, B. S. & Gertler, M. (1995), 'Inside the Black Box: The Credit Channel of Monetary Policy Transmission', *Journal of Economic Perspectives* **9**(4), 27–48.

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003), 'Latent Dirichlet Allocation', *Journal of Machine Learning Research* **3**, 993–1022.

Branch, W. A. (2004), 'The Theory of Rationally Heterogeneous Expectations: Evidence from Survey data on inflation expectations', *The Economic Journal* **114**(497), 592–621.

Bu, C., Rogers, J. & Wu, W. (2020), 'Forward-Looking Monetary Policy and the Transmission of Conventional Monetary Policy Shocks', *Finance and Economics Discussion Series 2020-014. Washington: Board of Governors of the Federal Reserve System* (2020-014).

Campbell, J., Evans, C., Fisher, J. & Justiniano, A. (2012), 'Macroeconomic Effects of Federal Reserve Forward Guidance', *Brookings Papers on Economic Activity* **43**(1), 1–80.

Carroll, C. D. (2001), 'The Epidemiology of Macroeconomic Expectations', *NBER Working Paper* (No. 8695).

Carroll, C. D. (2003), 'Macroeconomic Expectations of Households and Professional Forecasters', *Quarterly Journal of Economic* **118**(1), 269–298.

Coibion, O., Gorodnichenko, Y., Kumar, S. & Pedemante, M. (2020), 'inflation expectations as a policy tool?', *Journal of International Economics* (Forthcoming).

Cole, W. M. (2010), 'No News Is Good News: Human Rights Coverage in the American Print Media, 1980-2000', *Journal of Human Rights* **9**(3), 303–325.

D'Amico, S. & King, T. B. (2017), 'What Does Anticipated Monetary Policy Do?', *Federal Reserve Bank of Chicago* (WP 2015-10).

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.

Doms, M. & Morin, N. (2004), 'Consumer Sentiment, the Economy, and the News Media', *Finance and Economics Discussion Series Washington: Board of Governors of the Federal Reserve System* (2004-51).

Drager, L. & Lamla, M. J. (2017), 'Imperfect Information and Consumer Inflation Expectations: Evidence from Microdata', *Oxford Bulletin of Economics and Statistics* **79**(6), 933–968.

Durante, E., Ferrando, A. & Vermeulen, P. (2020), 'Monetary policy, investment and firm heterogeneity', *ECB Working Paper Series* (2390).

Dvorkin, M. A. & Shell, H. G. (2016), 'The Recent Evolution of U.S. Local Labor Markets', *Economic Synopses* (15).
URL: *https://research.stlouisfed.org/publications/economic-synopses/2016/08/01/the-recent-evolution-of-u-s-local-labor-markets/*

Federal Open Market Committee: Transcripts and other historical materials (2019).
URL: *https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm*

Federal Reserve Economic Data (2019).
URL: *https://fred.stlouisfed.org*

Feve, P. & Guay, A. (2016), 'Sentiments in SVARs', *Working Paper Tolouse School of Economics* (TSE 656).

Flack, E., Hoffman, M. & Hurtgen, P. (2019), 'Disagreement about inflation expectations and monetary policy transmission', *Journal of Monetary Economics* (Forthcoming).

Gertler, M. & Karadi, P. (2015), 'Monetary Policy Surprises, Credit Costs, and Economic Activity', *American Economic Journal: Macroeconomics* **7**(1), 44–76.

Gilbert, T., Scotti, C., Strasser, G. & Vega, C. (2016), 'Is the intrinsic value of macroeconomic news announcements related to their asset price impact?', *European Central Bank Working paper series* (1882).

Gilchrist, S. & Zakrajsek, E. (2012), 'Credit Spreads and Business Cycle Fluctuations', *The American Economic Review* **102**(4), 1692–1720.

Gomes, S., Iskrev, N. & Mendicino, C. (2017), 'Monetary policy shocks: We got news!', *Journal of Economic Dynamics and Control* **74**, 108–128.

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.

Gürkaynak, R. S., Kısacikoglu, B. & Wright, J. H. (2020), 'Missing Events in Event Studies: Identifying the Effects of Partially Measured News Surprises', *The American Economic Review* **110**(12), 3871–3912.

Gürkaynak, R. S., Sack, B. & Swanson, E. T. (2005), 'Do actions speak louder than words? the response of asset prices to monetary policy actions and statements', *International Journal of Central Banking* **1**(1), 55–93.

Gürkaynak, R. S., Sack, B. & Wright, J. H. (2007), 'The U.S. Treasury yield curve: 1961 to the present', *Journal of Monetary Economics* **54**(8), 2291–2304.

Gürkaynak, R. S., Sack, B. & Wright, J. H. (2010), 'The TIPS Yield Curve and Inflation Compensation', *American Economic Journal: Macroeconomics* **2**(1), 70–92.

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.

Hanson, S. G. & Stein, J. C. (2015), 'Monetary policy and long-term real rates', *Journal of Financial Economics* **115**(3), 429–448.

Henry, E. (2008), 'Are Investors Influenced By How Earnings Press Releases Are Written', *The Journal of Business Communication* **45**(4), 363–407.

Hirose, Y. & Kurozumi, T. (2012), 'Identifying News Shocks with Forecast Data', *CAMA Working Papers* (2012-01).

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.

Lamla, M. J. & Lein, S. M. (2014), 'The role of media for consumers' inftation expectation formation', *Journal of Economic Behavior and Organization* **106**, 62–77.

Larsen, V. H. & Thorsrud, L. A. (2019*a*), 'Business Cycle Narratives', *CESifo Working Paper* (7468).

Larsen, V. H., Thorsrud, L. & Zhulanova, J. (2020), 'News-driven inflation expectations and information rigidities', *Journal of Monetary Economics* (Forthcoming).

Larsen, V. & Thorsrud, L. (2019*b*), 'The Value of News for economic developments', *Journal of Econometrics* **210**(1), 203–218.

Le, Q. & Mikolov, T. (2014), 'Distributed Representations of Sentences and Documents', *Proceedings of the 31st International Conference on Machine Learning* **32**(2), 1188–1196.

Leduc, S. & Sill, K. (2013), 'Expectations and Economic Fluctuations: An Analysis Using Survey Data', *Review of economics and statistics* **95**(4), 1352–1367.

Leeper, E., Walker, T. & Yang, S.-C. (2013), 'Fiscal foresight and information flows', *Econometrica* **81**(3), 1115–1145.

Lei, C., Lu, Z. & Zhang, C. (2015), 'New on inflation and the epidemiology of inflation expectations in China', *Economic Systems* **39**(4), 644–653.

Levchenko, A. A. & Pandalai-Nayar, N. (2015), 'TFP, News, and "Sentiments": The International Transmission of Business Cycles', *NBER Working Paper* (21010).

Lewis, D. J., Makridis, C. & Mertens, K. (2020), 'Do Monetary Policy Announcements Shift Household Expectations?', *Federal Reserve Bank of New York Staff Report* (897).

Lorenzoni, G. (2009), 'A Theory of Demand Shocks', *The American Economic Review* **99**(5), 2050–84.

Lorenzoni, G. (2011), 'News and Aggregate Demand Shocks', *The Annual Review of Economics* **3**(1), 537–557.

Loughran, T. & Mcdonald, B. (2016), 'Textual Analysis in Accounting and Finance: A Survey', *Journal of Accounting Research* **54**(4), 1187–1230.

Lucas, R. (1972), 'Expectations and the neutrality of money', *Journal of Economic Theory* **4**(2), 103–124.

Lutkepohl, H. (2007), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag Berlin Heidelberg.

Mankiw, G. N. & Reis, R. (2002), 'Sticky Information versus sticky prices: a proposal to replace the New Keynesian Phillips curve', *Quarterly Journal of Economics* **117**(4), 1295–1328.

McCracken, M. W. & Ng, S. (2015), 'FRED-MD: A Monthly Database for Macroeconomic Research', *Federal Reserve Bank of St. Louis Working Paper* (2015-012B).

Michael D. Bauer and Eric T. Swanson (2020), 'The Fed's Response to Economic News Explains the "Fed Information Effect"', *NBER Working Paper* (27013).

Milani, F. & Treadwell, J. (2012), 'The Effects of Monetary Policy "News" and "Surprises"', *Money, credit and banking* **44**(8), 1667–1692.

Miranda-Agrippino, S. & Ricco, G. (forthcoming), 'The Transmission of Monetary Policy shocks', *American Economic Journal: Macroeconomics* .

Noh, E. (2018), 'Impulse-response analysis with proxy variables', *unpublished* .

Nowzohour, L. & Stracca, L. (2017), 'More than a feeling: confidence, uncertainty and macroeconomic fluctuations', *European Central Bank Working paper series* (No 2100).

Park, T. & Casella, G. (2008), 'The Bayesian Lasso', *Journal of the American Statistical Association* **103**(482), 681–686.

Plagborg-Mollerand, M. & Wolf, C. (2019), 'Local Projections and VARs Estimate the Same Impulse Responses', *Working paper: Department of Economics, Princeton University,* .

Ramey, V. (2016), 'Macroeconomic shocks and their propagation', *Handbook of macroeconomics* **2**, 71–162.

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.

Smith, A. L. & Becker, T. (2015), 'Has Forward Guidance Been Effective?', *Federal Reserve Bank of Kansas City Economic Review* **3**, 57–78.

Soo, C. (2015), 'Quantifying Animal Spirits: News Media and Sentiment in the Housing Market', *Ross School of Business Paper* (1200).

Souleles, N. S. (2004), 'Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys', *Journal of Money, Credit and Banking* **36**(1), 39–72.

Steinsson, E. N. J. (2019), 'Hight Frequency Identification of Money Non-Neutrality: the Information effect', *NBER Working Paper* (19260).

University of Michigan Survey of Consumers (2019).
URL: *https://data.sca.isr.umich.edu*

Woodford, M. (2001), 'Imperfect common knowledge and the effects of monetary policy', *NBER Working Paper* (8673).

Appendix A. News sentiments

? noted that for inflation and unemployment "an increase" denotes the bad state. The same intuition applies for the federal funds rate, the discount rate, and interest rates, which are here denoted as keywords together with inflation and unemployment. Therefore, if the following words appear near to the keywords they are labelled as negative, further to LM dictionary: negative, negatively, negatives, difficult, difficulty, hurdle, hurdles, obstacle, obstacles, uncertain, uncertainty, unsettled, unfavorable, depressed, disappoint, disappoints, disappointing, disappointed, disappointment, risk, risks, risky, threat, threats, penalty, penalties, deteriorate, deteriorates, deteriorating, deteriorated, worsen, worsens, worsening, worse, worst, challenge, challenges, challenging, challenged, up, increase, increases, increasing, increased, rise, rises, rising, rose, risen, exceed, exceeds, exceeded, exceeding, growth, up, high, higher, pessimism, more, above, high, higher, highest, greater, greatest, larger, largest, grow, grows, growing, grew, grown, growth, climbed.

The following words are labelled as positive if they appear near to the keywords: cut, cutback, cutbacks, deceased, decline, declined, declines, declining, diminish, diminished, diminishes, diminishing, downtime, downtimes, downturn, downturns, downward, downwards, dropped, neglect, neglected, neglectful, neglecting, neglects, negligence, negligences, negligent, negligently, shut, shutdown, shutdowns, shuts, shutting, slow, slowdown, slowdowns, slowed, slower, slowest, slowing, slowly, slowness, sluggish, sluggishly, sluggishness, weak, weaken, weakened, weakening, weakens, weaker, weakest, weakly, weakness, weaknesses, undercut, undercuts, undercutting, against, positive, positives, success, successes, successful, succeed, succeeds, succeeding, succeeded, accomplish, accomplishes, accomplishing, accomplished, accomplishment, accomplishments, strong, strength, strengths, certain, certainty, definite, solid, excellent, good, leading, achieve, achieves, achieved, achieving, achievement, achievements, progress, progressing, deliver, delivers, delivered, delivering, leader, leading, pleased, reward, rewards, rewarding, rewarded, opportunity, opportunities, enjoy, enjoys, enjoying, enjoyed, encouraged, encouraging, improve, improves, improving, improved, improvement, improvements, strengthen, strengthens, strengthening, strengthened, stronger, strongest, better, best, expand, expands, expanding, expanded, expansion, beat, beats, beating, fail, fails, failing, failure, weak, weakness, weaknesses, slump, slumps, slumping, slumped, downturn, down, decrease, decreases, decreasing, decreased, decline, declines, declining, declined, fall, falls, falling, fell, fallen, drop, drops, dropping, dropped, weaken, weakens, weakening, weakened, low, lower, lowest, less, least, cut, smaller, smallest, shrink, shrinks, shrinking, shrunk, below, under, deal, moderation, moderate, down, stop, stopping, deal, cool, optimism, stoppage, stoppages, stopped, stopping, stops, decline, lower, drop, decrease, slide.

I extract the five words that precede a keyword and the five words that follow it. If a sentence starts with a keyword I extract the seven following words, if a sentence ends with a keyword I extract the seven preceding words; in a short sentence I extract the words from the beginning of the sentence or the end of the sentence. After extracting words near a keyword I apply a lexical-based approach to label the sentiment of the keyword.

Additionally, I use a negation dictionary. If the following words precede a sentiment of keywords in the three-word window, then they are labelled as the opposite sentiment. The 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.

Appendix B. Latent Dirichlet Allocation

coherence.png

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

Table B.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, commun, 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 C. Cosine similarity and Euclidian distance

Cosine similarity between two vectors $\vec{A}$ and $\vec{B}$ is:

$$cos(\angle(\vec{A},\vec{B})) = \frac{\vec{A}\cdot\vec{B}}{||\vec{A}||_2||\vec{B}||_2}$$

where $\vec{A}\cdot\vec{B}$ is an inner product (a dot product in the Euclidean space) between two vectors and $||\vec{A}||_2||\vec{B}||_2$ is a product of their Euclidean lengths (L^2 norms).

$$||\vec{A}||_2 = \sqrt{\sum_i \vec{A}_i^{\,2}}$$

For unit-length vectors, cosine similarity is:

$$cos(\theta) = \vec{A}\cdot\vec{B}$$

since $||\vec{A}||_2 = ||\vec{B}||_2 = 1$. In this case minimisation of Euclidian distance (squared) is the same as maximisation of cosine similarity since:

$$Euclidian\_distance^2 = ||\vec{A}-\vec{B}||_2^2 = ||\vec{A}||_2^2 - 2\vec{A}\cdot\vec{B} + ||\vec{B}||_2^2 = 2 - 2cos(\angle(\vec{A},\vec{B}))$$

For an n-dimensional space the Euclidean distance is:

$$Euclidian\_distance = \sqrt{(a_1-b_1)^2 + (a_2-b_3)^2 + ... + (a_n-b_n)^2} =$$

$$= \sqrt{\sum_n (a_i-b_i)^2} = ||\vec{A}-\vec{B}||_2 = \sqrt{||\vec{A}||_2^2 + ||\vec{B}||_2^2 - 2\vec{A}\cdot\vec{B}}$$

Appendix D. Doc2Vec with k-means++

Table D.1: Topic labelling for Doc2Vec with kmeans++ 40 clusters

Topic	Words
new business	new, business, small, big, business, home, firm, sales, u.s., correction, appended, market, firms, industry, get, many, prices, still, help
dow	dow, stocks, shares, markets, bonds, market, rally, markets, trading, gains, prices, rise, wall, nasdaq, day, fall, investors, stock, earnings
jobs	job, workers, jobs, new, pay, firms, work, u.s., business, jobless, labor, rate, companies, may, economy, study, many, says, growth, california
profits	profit, earnings, company, quarter, loss, sales, reports, net, profits, posts, news, million, cut, business, rise, earnings, says, stock, earnings, jobs
housing	home, sales, housing, prices, real, mortgage, homes, rates, new, rise, market, price, starts, drop, newhome, fall, rate, u.s., may, county
reports	company, news, business, brief, rates, prices, earnings, profit, economy, rise, sales, briefing, reports, mortgage, u.s., briefing, rate, technology, jobless, fall
currency	dollar, currency, markets, gold, u.s., prices, yen, mixed, trading, markets, falls, gains, stocks, rises, mark, higher, new, lower, mostly, rise
fraud	case, fraud, says, million, sec, u.s., firm, suit, pay, former, court, stock, probe, accused, charges, judge, bank, new, may, firms
company stocks	stock, bid, company, buy, deal, firm, billion, offer, million, may, stake, news, shares, merger, new, sell, takeover, sale, business, market
farm prices	prices, food, u.s., farmers, farm, new, may, crop, industry, price, market, growers, sales, california, business, could, big, corn, says, demand
retailers	sales, retailers, stores, retail, new, holiday, walmart, online, shoppers, profit, chain, sears, retailer, store, may, big, company, shopping, market, business
energy	power, energy, gas, utility, new, u.s., california, electricity, utilities, state, plan, may, solar, natural, coal, plant, nuclear, california, business, electric
media	media, advertising, business, new, ad, advertising, ads, business, appended, correction, times, online, web, magazine, campaign, business, sales, tv, addenda, big
money	tax, money, personal, new, home, financial, money, may, retirement, loan, finance, mortgage, debt, college, savings, credit, loans, plan, student, help
international	world, u.s., debt, mexico, bank, new, economic, international, global, plan, latin, nations, aid, crisis, business, imf, economy, says, may, banks
economy	economic scene, tax, u.s., new, business, economy, market, correction, appended, may, economists, deficit, plan, budget, growth, james, economics, big, view
entertainment	company, town, tv, new, media, disney, sales, video, film, music, movie, may, deal, profit, hollywood, big, online, cable, digital, says
airlines	airlines, airline, air, travel, united, fares, business, fare, american, new, delta, cuts, u.s., fuel, company, may, cut, flights, loss, airways
financial markets	stocks, dow, market, markets, financial, rally, roundup, stock, investors, bond, wall, prices, nasdaq, oil, yields, gains, rise, tech, mixed, fall
banking	credit, card, bank, new, banks, cards, fees, consumer, personal, rates, consumers, online, may, money, pay, get, data, banking, customers, interest
economic	consumer, growth, u.s., sales, orders, economy, rise, rate, prices, spending, jobless, retail, economic, index, factory, inflation, economy, may, data, drop
deals	market, business, new, place, chief, correction, appended, people, wall, big, bank, company, u.s., stock, deal, executive, s.e.c., may, news, pay
services	phone, at&t, cable, fcc, new, wireless, company, service, tv, deal, internet, may, telecom, firm, plan, firms, mci, rates, bell, merger
oil/gas	oil, prices, gas, opec, price, u.s., gasoline, crude, output, rise, may, energy, cut, production, exxon, profit, company, new, pump, drop
real estate	real, estate, new, office, commercial, building, market, estate, city, downtown, may, housing, project, space, hotel, center, million, plans, plan, estate
loans	bank, banks, new, u.s., fed, loans, loan, mortgage, banking, big, profit, billion, financial, plan, says, s&l, credit, first, million, may
trading	trading, market, stock, futures, new, exchange, sec, nasdaq, wall, big, cbot, merc, nyse, board, options, markets, trade, chicago, may, plan
aircrafts	boeing, new, airbus, company, defense, u.s., air, orders, jet, deal, may, aircraft, business, says, billion, firm, lockheed, aerospace, pentagon, contract
vehicles	sales, auto, ford, car, gm, chrysler, u.s., new, g.m., company, cars, big, toyota, prices, profit, may, news, vehicle, says, plant
financial news	credit, markets, bond, prices, treasury, bonds, rates, yields, u.s., rise, issues, financenew, market, issues, treasuries, interest, new, fall, bonds, debt
health	drug, health, insurance, care, new, costs, price, may, medical, healthcare, business, u.s., says, insurers, prices, company, drugs, plan, firms, cost
business digest	business, digest, digest, week, economy, saturday, business, thursday, wednesday, friday, tuesday, monday, july, may, august
investing	market, funds, stocks, investors, mutual, stock, fund, place, wall, may, investing, new, beat, tom, money, bond, investing, investment, markets, funds
trade	trade, u.s., deficit, steel, talks, pact, imports, new, exports, may, says, global, gap, tariffs, world, mexico, foreign, trump, deal, house
fed	fed, rate, rates, interest, greenspan, inflation, says, economy, growth, economic, may, chief, fed's, cut, market, u.s., policy, money, bernanke, bank
cities	city, state, new, tax, chicago, plan, business, says, may, illinois, would, county, budget, could, d.c., economic, million, mayor, jobs
technology	new, apple, computer, company, sales, profit, technology, microsoft, market, chip, intel, pc, i.b.m., software, earnings, technology, news, business, ibm, stock
futures	prices, futures, futuresoptions, commodities, markets, oil, soybeans, grain, rise, corn, fall, wheat, coffee, gold, price, cattle, soybean, sharply, pork, drop
online	online, web, internet, new, google, technology, yahoo, business, firm, amazon, ad, company, microsoft, site, aol, tech, deal, firms, facebook, service
president	tax, obama, gop, house, plan, senate, bill, budget, new, bush, democrats, trump, u.s., president, says, debt, economic, may, deficit

Appendix E. Topic time series. Positive sentiments

corr1_nostation.jpg

Figure E.1: Cross-correlations between topic time series from the Doc2Vec model with k-means++

Consumer inflation, consumer interest and consumer unemployment are the expectations from the **?**

corr2_nostat.jpg

Figure E.2: Correlations between topic time series from the Doc2Vec model with k-means++ (y axis) and LDA using topic frequency labels (x axis)

Consumer inflation, consumer interest and consumer unemployment are expectations from the **?**

corr3_nostat.png

Figure E.3: Correlations between topic time series from the Doc2Vec model with k-means++ (y axis) and LDA using dominant topic labels (x axis)

Consumer inflation, consumer interest and consumer unemployment are expectations from the **?**

corr4_nostat.jpg

Figure E.4: Correlations between topic time series from LDA using dominant topic labels (y axis) and LDA using topic frequency labels (x axis) housing2.png

(a) Housing topics from Doc2Vec (red) and LDA (blue) using dominant topic labels

housing1.png

(b) Housing topics from Doc2Vec (red) and LDA (blue) using topic distributions labels

FED1.png

(c) FED topics from Doc2Vec (red) and LDA (blue) using dominant topic labels

FED2.png

(d) FED topics from Doc2Vec (red) and LDA (blue) using topic distributions labels

Figure E.5: Topics about Housing and Fed. All series are standardised.
Shaded areas - NBER based Recession Indicators for the United States

jobs1.png

(a) Jobs topics from Doc2Vec (red) and LDA (blue) using dominant topic labels

jobs2.png

(b) Jobs topics from Doc2Vec (red) and LDA (blue) using topic distributions labels

economic1.png

(c) Economic topics from Doc2Vec (red) and LDA (blue) using dominant topic labels

economic2.png

(d) Economic topics from Doc2Vec (red) and LDA (blue) using topic distributions labels

Figure E.6: Topics about Jobs and Economy. All series are standardised.
Shaded areas - NBER based Recession Indicators for the United States

currency1.png

(a) Currency topics from Doc2Vec (red) and LDA (blue) using dominant topic labels

currency2.png

(b) Currency topics from Doc2Vec (red) and LDA (blue) using topic distributions labels

oil:gas1.png

(c) Oil/gas topics from Doc2Vec (red) and LDA (blue) using dominant topic labels

oil:gas2.png

(d) Oil/gas topics from Doc2Vec (red) and LDA (blue) using topic distributions labels

Figure E.7: Topics about Currency and Oil/gas. All series are standardised.
Shaded areas - NBER based Recession Indicators for the United States

reports1.png

(a) Reports topics from Doc2Vec (red) and LDA (blue) using dominant topic labels

reports2.png

(b) Reports topics from Doc2Vec (red) and LDA (blue) using topic distributions labels

profits1.png

(c) Profits topics from Doc2Vec (red) and LDA (blue) using dominant topic labels

profits1.png

(d) Profits topics from Doc2Vec (red) and LDA (blue) using topic distributions labels

Figure E.8: about Reports and Profits. All series are standardised.

Shaded areas - NBER based Recession Indicators for the United States

loans1.png

(a) Loans topics from Doc2Vec (red) and LDA (blue) using dominant topic labels

loans2.png

(b) Loans topics from Doc2Vec (red) and LDA (blue) using topic distributions labels

Figure E.9: Topics about Loans. All series are standardised.
Shaded areas - NBER based Recession Indicators for the United States

hierarchical.png

Figure E.10: Hierarchical clustering of Doc2Vec topic time series using weighted linkage

Appendix F. LASSO results

Table F.1: Doc2Vec: LASSO results for interest rates

	positive	uncertainty	constraining	positive +AR(1)	uncertainty +AR(1)	constraining +AR(1)	positive× uncertainty+AR(1)	positive× constraining+AR(1)
new business								
dow								0.009
jobs	0.017			0.001				
profits	0.091			0.02			0.004	
housing				0.006				
reports	0.092			0.027				
currency								
fraud								
company stocks	0.01			0.002				
farm prices								
retailers								
energy								
media						-0.014		
money	0.01							
international							0.006	
economy								0.001
entertainment								
airlines								
financial markets				0.011				
banking								
economic	0.014							
deals								
services								
oil/gas								
real estate								
loans								0.002
trading							0.006	
aircrafts								
vehicles								
financial news								
health								
business digest								
investing				0.001				
trade								
fed				0.014			0.006	0.002
cities								
technology	0.028			0.006				0.001
futures								
online				0.002				
president								

Table F.2: Doc2Vec: LASSO results for unemployment

	positive	uncertainty	constraining	positive +AR(1)	uncertainty +AR(1)	constraining +AR(1)	positive× uncertainty+AR(1)	positive× constraining+AR(1)
new business								
dow			0.005					-0.024
jobs			0.002		0.004			
profits			-0.013					
housing	-0.104		0.012	-0.005				
reports								
currency			0.002					
fraud			-0.01					
company stocks			0.038		0.012			
farm prices			-0.009					
retailers								
energy								
media			0.063					
money		0.089	0.137		0.002			
international			-0.065					
economy								
entertainment			-0.004					
airlines			0.068					
financial markets	-0.157		-0.034	-0.056				
banking			-0.001					
economic			-0.013					
deals			-0.035					
services							0.003	0.007
oil/gas								
real estate			-0.013					
loans								
trading								
aircrafts								
vehicles		0.004						
financial news			-0.028			-0.005	-0.016	
health			0.014					
business digest			-0.078					
investing								
trade								
fed								
cities		-0.149			-0.029			
technology			-0.008					
futures			0.009					
online				-0.015				
president			0.027	-0.002	0.006	0.002		

Table F.3: Doc2Vec: LASSO results for inflation

	positive	uncertainty	constraining	positive +AR(1)	uncertainty +AR(1)	constraining +AR(1)	positive× uncertainty+AR(1)	positive× constraining+AR(1)
new business		0.015	0.016	0.007				
dow		-0.033	0.034		-0.008		0.002	
jobs		0.04			0.006		-0.022	
profits		-0.013	-0.046					
housing	-0.07			-0.009				
reports	0.001	0.019	0.03					-0.011
currency		0.029	0.007					
fraud			0.011					
company stocks		0.017						
farm prices		0.005			0.001			
retailers		-0.044	-0.075			-0.002		
energy	0.005		-0.016					
media								
money			0.027					-0.001
international		0.002						
economy		-0.017	-0.076					
entertainment								
airlines		0.071	0.056	0.011			0.012	0.015
financial markets	-0.029							
banking		-0.094	-0.071	0.002				
economic		0.01	0.001					
deals			-0.018					
services			0.008					
oil/gas		0.19	0.161	0.003	0.009	0.003		
real estate			-0.003	-0.002			-0.002	
loans	-0.217	-0.013	0.046	-0.016		0.001		
trading		-0.075	-0.02					
aircrafts		-0.008	0.019					
vehicles		0.032						
financial news			-0.021					
health				0.001	-0.001		0.002	
business digest		-0.076	-0.104					
investing		-0.011						
trade								
fed		-0.038						
cities		-0.199	-0.02					
technology		-0.006	-0.083			-0.001		
futures	-0.056		0.116		0.012	0.02		-0.015
online		-0.003						
president		-0.009						

Table F.4: LDA using topic frequency labels: LASSO results for interest rates

	positive	uncertainty	constraining	positive +AR(1)	uncertainty +AR(1)	constraining +AR(1)	positive× uncertainty+AR(1)	positive× constraining+AR(1)
rates								
computers	0.012			0.005			0.007	
economic	0.002	-0.188	-0.107	0.076	-0.06	-0.017	0.069	0.066
food								
people								
media								
fed							0.003	
housing								
credit								
cars								
health								
trade								
law								
debt								
loans								
stocks								0.005
schools								
economics					0.001			
retailers								
industry	0.095			0.007				
cities								
profits	0.072			0.018				
jobs	0.129							
currency								
airlines								
military								
energy								
oil/gas	0.043	0.011						
international								
hotels								
rules								
stock market								
company news								
services								
investing								
president								
reports								
securities								
budget								
deals							0.004	

Table F.5: LDA using topic frequency labels: LASSO results for unemployment

	positive	uncertainty	constraining	positive +AR(1)	uncertainty +AR(1)	constraining +AR(1)	positive× uncertainty+AR(1)	positive× constraining+AR(1)
rates					-0.001			
computers								
economic	-0.234	0.13						
food								
people								
media								
fed							-0.008	
housing	-0.157	0.27	0.028	-0.026	0.07		-0.038	-0.051
credit								
cars								
health								
trade								
law								
debt								
loans								
stocks				-0.059				
schools								
economics								
retailers					0.004			
industry								
cities								
profits					-0.005			
jobs					0.014		-0.024	-0.001
currency								
airlines								
military					-0.019			
energy								
oil/gas		0.016						
international								
hotels								
rules					0.007		-0.015	-0.004
stock market		0.008					-0.001	
company news								
services								
investing								-0.01
president				-0.016				
reports								
securities								
budget								
deals					0.003			

Table F.6: LDA using topic frequency labels: LASSO results for inflation

	positive	uncertainty	constraining	positive +AR(1)	uncertainty +AR(1)	constraining +AR(1)	positive× uncertainty+AR(1)	positive× constraining+AR(1)
rates					-0.001			0.017
computers								
economic		-0.037	-0.007					0.034
food		0.178	0.124		0.04	0.02		
people								
media								
fed	-0.002	-0.029			-0.007			
housing	-0.006	0.029					-0.011	-0.023
credit	0.087	-0.044	-0.021	0.03			0.005	0.026
cars		-0.031			-0.009			
health			-0.01	0.002	-0.013	-0.023	0.02	0.025
trade		0.014			0.006			
law				-0.007		0.003		-0.005
debt								
loans	-0.452	0.148	0.228	-0.078	0.021	0.038	-0.004	-0.058
stocks	-0.002	-0.019			-0.003			0.018
schools					0.003			0.005
economics			-0.025					
retailers		-0.037	-0.009					
industry	0.03	-0.005						
cities		-0.035	-0.072			-0.01		
profits	0.027	-0.015						
jobs			-0.008		0.007			-0.005
currency					0.009	0.024		0.003
airlines	0.028			0.019	-0.012	-0.011	0.05	0.038
military		-0.184	-0.145	0.012	-0.046	-0.025		0.007
energy								
oil/gas		0.322	0.207	0.019	0.033	0.01		
international					0.001	0.008		
hotels		0.007				0.004		
rules								
stock market								
company news								
services		-0.007			-0.006			-0.002
investing								
president		-0.035			-0.014	-0.004		
reports								0.001
securities						0.001		
budget			0.004			0.01		-0.013
deals								-0.005

Table F.7: LDA using dominant topic labels: LASSO results for interest rates

	positive	uncertainty	constraining	positive +AR(1)	uncertainty +AR(1)	constraining +AR(1)	positive× uncertainty+AR(1)	positive× constraining+AR(1)
rates								
computers								
economic	0.052	-0.067	-0.053	0.053	-0.041	-0.035	0.039	0.047
food								
people								
media	0.008							
fed				0.009			0.009	0.012
housing								
credit								
cars								
health								
trade								
law								
debt								
loans						-0.005		0.007
stocks	0.004			0.011				
schools								
economics								
retailers								
industry	0.048							
cities								
profits	0.123			0.037			0.023	0.027
jobs	0.007							
currency								
airlines								
military								
energy								
oil/gas								
international								
hotels								
rules								
stock market								
company news				0.015				
services								
investing								0.001
president								
reports								
securities							0.001	
budget								
deals				0.002				

Table F.8: LDA using dominant topic labels: LASSO results for unemployment

	positive	uncertainty	constraining	positive +AR(1)	uncertainty +AR(1)	constraining +AR(1)	positive× uncertainty+AR(1)	positive× constraining+AR(1)
rates								
computers								
economic	-0.109		0.009					
food								
people								
media								
fed								
housing			0.006			0.012		
credit						-0.001		
cars	-0.007							
health								
trade								
law								
debt								
loans	-0.049		0.012	-0.005		0.004		-0.009
stocks	-0.039			-0.048				
schools								
economics								
retailers						0.002		
industry								
cities								
profits								
jobs								
currency								
airlines								
military							0.021	0.023
energy								
oil/gas								
international			-0.027			-0.015		
hotels						-0.011		
rules								
stock market								
company news								
services								
investing								
president				-0.012		0.013		
reports								
securities					-0.003	-0.031	-0.017	
budget								
deals								

Table F.9: LDA using dominant topic labels: LASSO results for inflation

	positive	uncertainty	constraining	positive +AR(1)	uncertainty +AR(1)	constraining +AR(1)	positive× uncertainty+AR(1)	positive× constraining+AR(1)
rates								
computers							0.003	
economic		-0.03		0.008			0.027	0.01
food		0.129	0.098		0.009	0.009		
people								
media								
fed							0.005	
housing			0.043			0.001	-0.002	
credit		-0.009		0.003			0.012	
cars							0.008	0.007
health				0.005		-0.007	0.008	
trade								
law							-0.003	
debt								
loans	-0.282		0.207	-0.036		0.032		-0.01
stocks								
schools								
economics			-0.01					
retailers								
industry	0.006							
cities				-0.002				
profits								
jobs								
currency			0.067			0.019		-0.01
airlines	0.031			0.022			0.028	0.037
military							0.003	
energy								
oil/gas		0.189	0.11					
international								
hotels			-0.054					
rules								
stock market								
company news								
services								
investing							-0.008	
president								
reports								
securities								
budget								
deals								

factor_loadings.png

Figure F.1: Factor loadings

correl.png

Figure F.2: Correlations between the first principal component from news and the first factor from the FRED-MD (McCracken & Ng (2015)) at leads and lags

Negative numbers are lags, positive numbers are leads

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):

$$p(\beta, Q) = p(\beta)p(Q)$$
$$p(\beta) \sim f_N(\beta|\underline{\beta}, \underline{V_\beta})$$
$$p(Q) \sim f_{IW}(Q|\underline{Q}, \underline{v_Q})$$

To deal with overfitting I entertain a prior in Minnesota fashion. Prior for β is set at its univariate AR(p) estimate, 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: $\lambda_1 = 0.2$, $\lambda_2 = 0.5$, $\lambda_3 = 1$, $\lambda_4 = 100$. σ_i, σ_j are scaled measures of the variance associated with the AR(p) equation estimate. $\underline{Q}$ is a diagonal matrix with diagonal elements equal to its initial OLS estimate. Lastly, I set $\underline{v_Q} = 30$. Based on the priors the conditional posterior for β is:

$$\beta|y, Q^{-1} \sim N(\overline{\beta}, \overline{V_b})_{I_{s(\beta)}}$$
$$\overline{V_\beta} = (\underline{V_\beta}^{-1} + \sum_{t=1}^{T} X_t' Q^{-1} X_t)^{-1}$$
$$\overline{V_b} = \overline{V_\beta}(\underline{V_\beta}^{-1}\underline{\beta} + \sum_{t=1}^{T} X_t' Q^{-1} y_t)$$

$I_{s(\beta)}$ is an indicator function used to denote that the roots of β lie outside the unit circle.

The conditional posterior of Q is:

$$Q|y, \beta \sim IW(\overline{Q}, \overline{v_Q})$$
$$\overline{v_Q} = \underline{v_Q} + T$$
$$\overline{Q} = \underline{Q} + \sum_{t=1}^{T} (y_t - X_t'\beta)(y_t - X_t'\beta)'$$

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

Appendix H. Soft news and real activity

FEVD_front.png

(a) Soft news shock. Ordered first

FEVD_back.png

(b) Soft news shock. Ordered last

Figure H.1: Contributions of shocks to forecast error variances. SVAR using the principal component from news

Without the consumer sentiment index

The numbers are based on the median impulse response functions

FEVD_soft.png

(a) Soft news shock. Ordered last

FEVD_centiment.png

(b) Consumer sentiment shock. Ordered second last

Figure H.2: Contributions of shocks to forecast error variances. SVAR using the principal component from news and consumer sentiments

Using the consumer sentiment index

The numbers are based on the median impulse response functions

real.jpg

(a) IRFs to a soft news shock. Ordered last

FDPE.png

(b) Soft news shock for forecast error variances

Figure H.3: Soft news shock. SVAR using the principal component from news
The real rate is the FFR less expected inflation

Appendix I. Heterogeneity of shocks

unemp_exp.jpg

(a) Unemployment sentiment shock, ordered second last

inflation2.jpg

(b) Inflation sentiment shock with loans, ordered second last

exp.jpg

(c) Interest rate sentiment shock with economic topic, ordered second last

Figure I.1: Sentiment shocks. SVARs using news topics and expectations
median and 16th and 84th percentiles

HOUSING.png

(a) Housing news shock

unemp_exp.png

(b) Unemployment sentiment shock

Figure I.2: Contributions of shocks to forecast error variances. SVAR using unemployment expectations and the Housing topic

The numbers are based on the median impulse response functions

loans.png

(a) Loans news shock

inf_exp.png

(b) Inflation sentiment shock

Figure I.3: Contributions of shocks to forecast error variances. SVAR using inflation expectations and the Loans topic

The numbers are based on the median impulse response functions

economic.png

(a) Economic news shock

interest_exp.png

(b) Interest rate sentiment shock

Figure I.4: Contributions of shocks to forecast error variances. SVAR using interest rate expectations and the Economic topic

The numbers are based on the median impulse response functions

oil.jpg

(a) Oil/gas news shock, ordered last

inf_exp.jpg

(b) Inflation sentiment shock, ordered second last

Figure I.5: Soft news and sentiment shocks. SVAR using inflation expectations and the Oil/gas topic

median and 16th and 84th percentiles

oil_gas.png

(a) Oil/gas news shock

inf_exp2.png

(b) Inflation sentiment shock

Figure I.6: Contributions of shocks to forecast error variances. SVAR using inflation expectations and the Oil/gas topic

The numbers are based on the median impulse response functions

Appendix J. Soft news and monetary policy

1ipNOU.jpg

(a) Impulse responses to a Fed news shock (from Doc2Vec)

2ip.jpg

(b) Impulse responses to a Fed news shock (sentiment uncertainty)

2ipNOU.jpg

(c) Impulse responses to Fed news shocks

Figure J.1: Impulse responses to Fed news shocks from the LDA model using topic frequency labels

median and 16th and 84th percentiles

1.jpg

(a) Impulse responses using CFNAI instead of industrial production

2.jpg

(b) Impulse responses using CFNAI (from the LDA frequency model)

Figure J.2: Impulse responses to a Fed news shock using sentiment uncertainty
median and 16th and 84th percentiles

2008_2014_l6.jpg

Figure J.3: Impulse responses to a Fed news shock, 2008:M1–2014:M12, 6 lags

median and 16th and 84th percentiles

ec22.png

(a) Economic news shock

fed22.png

(b) Fed news shock

Figure J.4: Contributions of shocks to forecast error variances. SVAR using the Economic and the Fed topics

The numbers are based on the median impulse response functions

loans4.png

(a) Loans news shock

Figure J.5: Contributions of shocks to forecast error variances. SVAR using the Economic and the Loans topics

The numbers are based on the median impulse response functions

loans5.png

(a) Loans news shock

expect2.png

(b) Sixth shock

Figure J.6: Contributions of shocks to forecast error variances. SVAR using the Economic, the Loans topics and inflation expectations

The numbers are based on the median impulse response functions

Appendix K. Information augmentation for conventional monetary policy shocks

A conventional problem in identifying monetary policy shocks is that the information set of a decision-maker, the Federal Reserve in this case, is larger than the information set of the econometrician, which is the variables included in an econometric model. In this case a shock from the econometric model is not correctly identified, which is the well-known problem of nonfundamentalness[50].

Given this, **?** augmented the standard monetary VAR with the first principal component from the FRED-MD database to take unobservables about economic conditions into account. Therefore, I additionally augment the model with the first principal component from the news topic time series to take the information set of private agents into account as well. Besides sentiment, the news media might capture unobserved fundamentals or unobserved information.

The variables in the VAR are the logarithm of industrial production (IPB50001N), the logarithm of the consumer price index (CPIAUCNS), and the federal funds rate (FEDFUNDS). All the variables are taken from the *Federal Reserve Economic Data* (2019) and the period studied is 1984:M1–2019:M7. The impulse responses of real economic activity to a monetary policy shock are presented in Figure K.1[51]. The VAR estimation details are presented in Appendix G. Identification is achieved via standard recursive ordering: industrial production, inflation, the federal funds rate, the first factor from the FRED-MD, and the first principal component from the topic time series. The results are robust to re-ordering of the informational variables.

[50] This problem was pointed out by **?**, **?**, **?**, and Leeper et al. (2013) among others.

[51] The impulse responses of inflation show a price puzzle of similar magnitude in all specifications.

3var.png

(a) 3 variable SVAR

factor.png

(b) 3 variable SVAR using FRED-MD

pca.png

(c) 3 variable SVAR using news sentiments

both.png

(d) 3 variable SVAR using FRED-MD and news sentiments

Figure K.1: Impulse responses of industrial production to a monetary policy shock using additional information variables

median and 16th and 84th percentiles