What Say They About Their Mandate? A Textual Assessment of Federal Reserve Speeches

Myrthe van Dieijen[1] and Robin L. Lumsdaine[1,2]
August 14, 2019

Abstract

This paper considers the extent to which the Federal Reserve has balanced the responsibilities set out by its dual mandate and whether, in the aftermath of the 2007-8 financial crisis, the Federal Reserve effectively adopted a third mandate of ensuring financial stability and mitigating systemic risk. We do this by analyzing officials' speeches using a variety of text-processing techniques. We find that the price stability emphasis that existed before the crisis gave way to a focus on employment conditions as the crisis unfolded. Additionally, we indeed find an increasing role for financial stability in the post-crisis period.

Keywords: Financial Stability; Dual Mandate; FOMC; Latent Dirichlet Allocation (LDA); Dynamic Topic Model (DTM)
JEL Classification Codes: E58, E61, C54, C63, E02

[1] Erasmus School of Economics, Erasmus University Rotterdam
[2] Kogod School of Business, American University; National Bureau of Economic Research; Center for Financial Stability; Tinbergen Institute

Acknowledgements: Joseph Shadel provided excellent research assistance in collecting some of the speeches used in this paper. Financial support for the first author from the Erasmus School of Economics and the Erasmus Research Institute of Management during her PhD is gratefully acknowledged. Lumsdaine is grateful for the generous hospitality extended during various visits to both Erasmus University Rotterdam and the University of Portsmouth in facilitating the collaboration on this project.

The views and opinions expressed are those of the authors only and do not necessarily represent the views of any institutions with which they are affiliated.

'The success of monetary policy should be judged by the economy's performance against our statutory mandates of price stability and maximum employment.'

--

Jerome H. Powell

'Thoughts on the Normalization of Monetary Policy'

Speech delivered at the Economic Club of New York, New York

June 1, 2017

1. Introduction

For nearly four decades, the Federal Reserve ("Fed") has operated under what is commonly referred to as a "dual mandate", loosely defined as a responsibility to balance the objectives of stable prices with sustained growth. It is therefore easy to forget that prior to 1977, when Section 2A was added to the Federal Reserve Act, no explicit mandate was specified, although some have argued that a single mandate was implicit.[1] And despite the "dual mandate" moniker, the actual language of the 1977 Amendment identifies "goals of maximum employment, stable prices, and moderate long-term interest rates". Although Williams (2012) clearly refers to "the Fed's three mandates", this is unusual among Federal Reserve officials. For example, the July 26, 2017 FOMC statement noted that, "Consistent with its statutory mandate, the Committee seeks to foster maximum employment and price stability", without any mention of a third goal. In fact, this exact sentence has been a fixture in all statements issued in conjunction with scheduled FOMC meetings since November 3, 2010 (a total of 73 statements,

[1] Haltom and Lacker (2013) argue that it was monetary stability while Kahn and Taylor (2014) suggest it was financial stability. In addition, Steelman (2011) notes that even after the 1977 Amendment, there were episodes when the Fed was perceived as ignoring the employment part of the dual mandate and focusing exclusively on price stability.

as of July 31, 2019), indicating that this particular articulation of the dual mandate has been fairly codified in official FOMC statements for quite some time.

In their speeches, however, Federal Reserve officials have at times deviated from this "official" definition, likely because a speech provides a medium through which additional words of explanation or clarification can be offered. In speeches, for example, the mandate commonly has been articulated (see, e.g., Pianalto 2011, Williams 2012) as balancing the goals of price stability and maximum *sustainable* employment (emphasis added; Kahn and Taylor 2014 credit the Greenspan era with the introduction of this emphasis while Steelman 2011 cites November 2010 as the point of introduction), in recognition of the fact that a literal interpretation of "maximum employment" would mean that every individual is working at all hours of the day and night (Williams 2012). Both the evolution and fluidity of the interpretation of the mandate, relative to the terminology given in the Federal Reserve Act, have been discussed by Steelman (2011) and Kahn and Taylor (2014). These articles demonstrate that while the actual mandate (as written in the Federal Reserve Act) has not changed in nearly 40 years, the articulated mandate (as provided in Federal Reserve officials' speeches and press reports) has varied over time.

Beyond the goals of the mandate, however, is the question of how these goals should be weighted in policy decisions. For example, in 2007, then-Chairman Ben Bernanke explicitly referred to an equal weighting between the two components of the dual mandate (Bernanke 2007, as cited in Kahn and Taylor 2014). Yet as noted in Kahn and Taylor (2014), "Interpretation of the dual mandate has varied over time depending on economic circumstances and policymakers' understanding of the economy." In this paper, we are interested in the extent to which shifts in emphasis can be detected in Federal Reserve speeches. The reason for this

interest is twofold: (1) market participants frequently look to Federal Reserve speeches for evidence of the current and future policy stance, (2) policymakers may want to identify situations where the perceptions being given by their communications do not align with their intended policy stance.

The "unusual and exigent circumstances"[2] that surrounded the recent global financial crisis and led to a sustained period of quantitative easing has led to increased scrutiny as to the specifics of the Fed's mandate. While some have made reference to the single price stability mandate of the ECB (e.g., Bullard 2010), most of the post-crisis discussion has centered on the extent to which financial stability should be added to, or already is part of, the Fed's mandate. It is evident that even among FOMC members themselves, there is no clear consensus. Some FOMC members (and coauthors) have opined on whether the mandate should be altered to include financial stability (Haltom & Lacker, 2013, Peek, Rosengren, and Tootell 2016, Mester 2016), while others (Brainard 2014 and Dudley 2016) argue it has already been added (this latter view is echoed in Baxter 2013 and Michel 2014). In light of this lack of consensus, a second area of interest in this paper is whether the Fed's mandate has changed to include financial stability *de facto*, even if not *de jure.* Identification of such changes has perhaps become more important as central banks provide guidance as they exit from the prolonged period of hovering on- or near the zero-lower bound.

[2] This is a reference to the language in Section 13(3) of the Federal Reserve Act that was first invoked in March 2008 in the context of the Bear Stearns-JP Morgan Chase arrangement and subsequently used repeatedly to justify a series of programs undertaken by the Fed to address the financial crisis. Interestingly, while Section 2A delineated objectives for "The Board of Governors of the Federal Reserve System and the Federal Open Market Committee", in Section 13(3) powers are extended solely to the Board of Governors.

To examine these two questions of interest, namely (a) whether there has been a shift in the weights placed on the dual mandate goals of maximum employment and stable prices, and (b) whether there is evidence that financial stability has become a *de facto* third goal of the Fed's mandate, a variety of techniques from the field of text analysis are applied to speeches made by members of the Board of Governors of the Federal Reserve System from 1997-2016.

The next section provides a brief summary of previous literature examining Federal Reserve communications, particularly those works that have employed text analysis. Section 3 describes the speech data used in the analysis. Section 4 summarizes the methodologies used to analyze the data and presents descriptive results. Section 5 provides an econometric analysis of the main question of interest, the extent to which the weights associated with the three potential components of the Fed's mandate have changed over time. Section 6 concludes.

2. Brief Review of Text Literature

Several studies have investigated various types of Fed communication[3] in relation to market and/or economic variables (Kohn and Sack, 2004; Lucca and Trebbi, 2009; Hansen and McMahon, 2016), or the federal funds rate (Hansen, McMahon, and Prat, 2014; Peek, Rosengren and Tootell, 2016) by employing natural language processing techniques. Other papers have analyzed meeting transcripts using natural language processing to explore the deliberation process of the FOMC committee (Bailey and Schonhardt-Bailey, 2008; Fligstein, Brundage, and Schultz, 2014; Hansen, McMahon and Pratt, 2014; Schonhardt-Bailey, 2013), or to study rhetoric changes (Abe, 2011; Bligh and Hess, 2007).

[3] e.g., FOMC meeting transcripts, minutes, statements and speeches by Fed governors.

More recently, content analysis has been applied to FOMC meeting transcripts to (a) consider the possibility that financial stability should be adopted as a third mandate (Peek et al., 2016) and (b) identify whether financial stability already seems to matter in terms of monetary policy decisions (Oet and Lyytinen 2017). This earlier work has sought to identify the impact on monetary policy by evaluating whether the addition of text-generated measures of financial (in)stability can help to explain deviations from the Taylor rule. Part of the challenge in considering such an impact is the definition and identification of "financial stability". Peek et al. (2016) measure the intensity of the financial instability concerns by computing a simple word count of terms thought to be related, that is, by means of a "self-computed dictionary". They then include this intensity measure in a Taylor rule specification that includes forecasts of inflation and unemployment but does not include analogous word counts for the two existing mandates. It is difficult to draw inference about the Fed's potential focus on financial instability, therefore, as the number or words related to financial stability might still be negligible compared to word counts concerning the two existing mandates. This drawback is addressed in Oet and Lyytinen (2017), who extract thematic variables as measures for financial stability and the two mandates and show that these themes vary over time.[4] The paper goes on to demonstrate that adding these themes improves the fit of Taylor-rule and Taylor-rule-type models.

Beyond the question of whether financial stability has officially become part of the mandate and influences the FOMC's monetary policy decisions is the question of whether the Fed's communications have signaled a change in focus and if so, when that change occurred.

[4] More precisely, Oet et al. (2017) use six thematic variables: inflation, output, unemployment, financial stability, foreign activity, and fiscal policy.

Thus in contrast to Peek et al. (2016) and Oet and Lyytinen (2017), our paper aims to identify whether such a shift has occurred rhetorically -- with financial stability being a *de facto* third mandate, in addition to the existing two mandates. Moreover, we investigate the impact of Fed *speeches*, rather than FOMC meeting transcripts, as the former are available in real time and vetted extensively.[5] In our view, speeches are a more likely vehicle for shaping external constituents' current thinking than transcripts because transcripts are released with some lag and thus are unlikely to be a primary vehicle for communicating and shaping a message to influence public perception. In addition, because transcripts are a record of the actual monetary policy discussions, they reflect aggregate (rather than individual) thinking and hence might not capture individual heterogeneity regarding the balance of weights across the (two- or three-) mandates.[6]

The text analysis methods we employ map the change in the Fed's perception of its responsibilities over time. To quantify this shift we use advanced topic models and compare the results to the inference drawn from analysis employing simple word counts or dictionaries. We start our explanation of the topic models with a frequently-used more advanced model, the Latent Dirichlet Allocation (LDA) model (Blei et al., 2003), which has been shown to be one of the most popular models for eliciting underlying topics of a text. The model is equipped to deal

[5] Fed speeches are typically released early on an embargo basis (i.e., with the requirement that they not be released or referred to in advance of their delivery) in order that media outlets can prepare their reports and comments in advance, for release to coincide with the speech being delivered. In addition, the speech is usually released on the Fed's website within minutes of the start of its delivery.

[6] Of course, to address this criticism, one could potentially conduct a content analysis of the subset of text related to each individual speaker in the transcripts but to our knowledge, this has not been done.

with large collections of texts and has served as a basis for a variety of other topic models (Ramage, Dumais and Liebling, 2010; Srivastava and Sahami, 2009).[7]

One of the limitations of the LDA approach, however, is that it is static and can thus only be used to identify topics of separate time slices without assuming those topics are connected over time. Because we are interested in how the Fed has balanced its particular mandates and whether this balance has changed over time, we additionally employ the Dynamic Topic Model (Blei et al., 2006) for our analysis. This model is a dynamic adaptation of the LDA model and is well equipped to deal with time-varying topics, which is why its results will be the focal point of our analysis. To our knowledge, we are the first to use this model in the context of the Fed's mandates.

Much of the aforementioned literature analyzes FOMC meeting transcripts or minutes. In research where the content of Fed governors' *speeches* are investigated (Kohn and Sack, 2004; Bligh and Hess, 2007; Abe, 2011), the analysis has been limited to speeches of Alan Greenspan. In contrast, we investigate the speeches of all the Fed governors, as those texts are carefully constructed and poured over by staff in advance of their delivery. By examining the speeches of different governors, we are able to capture potential heterogeneity or disagreement regarding what should be the Fed's focus. To our knowledge, this paper is the only one to study speeches from multiple governors via natural language processing.

[7] The use of the LDA model (and its extensions) has not been limited to text data, as it also has been successfully applied to computer images (Fei-Fei and Perona, 2005), survey data (Erosheva, Fienberg and Joutard, 2007), and social network data (Jiang, Qian, Shen, Fu, and Mei 2015) for example.

3. Data

We downloaded 1229 speeches delivered by members of the Board of Governors of the Federal Reserve System between 1997 and 2016 inclusive.[8] For instances where the same remarks were delivered on more than one occasion (as noted by the Federal Reserve's website), only the first instance of the speech was included (to avoid double counting).[9] In addition, speeches characterized as commencement addresses, opening/welcoming/closing remarks or unveilings or dedications of some sort were deleted, amounting to a total of 56 deleted speeches.[10] Hence, the final sample consists of 1173 speeches. After lower casing all words and

[8] These are available at http://www.federalreserve.gov/newsevents/speech/YYYYspeech.htm, for each year YYYY = 1996…2017. Because when we began our analysis, 1996 and 2017 were only partial years, we restricted our sample to the complete years in the database (1997…2016). The texts were downloaded as pdf files, and then converted to txt format. These texts were then cleared of appendices, footnotes, figures, and tables. In addition, the texts were cleared of hyperlinks, page numbers, and headers that marked the time zones, embargo instructions (e.g., 'For release at'), and the delivery format of the speech (e.g., 'via videoconference'). Headers or footers that resulted from the pdf download (e.g., 'Printer Version', or 'Return to top') were also deleted. The resulting collection of texts all start with the name of the governor (e.g., 'Remarks by Janet Yellen', or 'Speech by Alan Greenspan), followed by the title, location and date (in varying order), and the speech itself.

[9] We additionally deleted the March 25, 1999 speech by Edward Kelley and the October 6, 1999 speech by Roger Ferguson, because they were very similar to the ones given on October 29, 1998 and September 29, 1999, respectively, despite not being noted as the same on the website. Moreover, the speech on March 23, 2003 by Donald Kohn was deleted because, despite being listed on the website as a speech, it was a journal article written in cooperation with Brain P. Sack.

[10] More precisely, 46 speeches were characterized as commencement addresses, welcoming/opening/closing remarks, dedications (e.g., the March 7, 2009 remarks given by Chairman Bernanke at the Interstate Interchange Dedication Ceremony naming an interstate exit in his honor), or concerned other types of ceremonial activities (e.g.,

clearing the texts from punctuation and (numerical) symbols, this sample contains 3,788,490 words (*tokens*) in total, of which 24,116 are unique words (*types*).

3.1 Description of the speeches

An overview of the Fed governors in our sample, along with their position (i.e., where a Governor later became a Vice-Chair or Chair, we note these episodes separately), date of term, number of speeches, tokens (all words) and types (unique words) per position, is presented in Table 1. Not only do longer speeches contain more tokens, they also are more likely to have a higher number of types; therefore, an often-used metric is the type-to-token ratio (i.e., the number of types in a speech divided by the number of tokens). It serves as a measure of lexical diversity, as it indicates what percentage of the total number of words (tokens) is unique. A higher type-to-token ratio indicates greater lexical diversity. The table thus includes the average type-to-token ratio for each individual calculated separately over the speeches given while holding a particular position (e.g., the average TTR of the speeches Ben Bernanke gave while governor is 0.27 and while chair is 0.32).

[Insert Table 1 here]

Figure 1 displays the number of speeches that were given in each year of the sample period. The number of speeches per year increased in the start of our sample, reaching a maximum of 95 in 2004. Hereafter, the number of speeches decreased, with a distinct drop

---

the unveiling of a new $100 note). In addition, seven speeches were deleted because they concern the Economic Growth and Regulatory Paperwork Reduction Act (EGRPA). The other three omitted files are the ones mentioned in footnote 8.

noticeable in 2009. This drop is partially explained by the lower number of serving governors since then, as shown by considering the number of types and tokens per speech, displayed in Figure 2.

[Insert Figure 1 here]

[Insert Figure 2 here]

The relationship between the number of types and tokens of the speeches in our sample can be seen in Figure 3. Not surprisingly, there is a high correlation between the number of types and tokens per speech (0.923). The cluster of points reveals that for most of the speeches the number of types and tokens varies between 500 and 1500, and 2000 and 6000, respectively. Not all speeches lie within those ranges, however; specifically, there are a few speeches (in the top right corner of the plot) that have a very high number of tokens. The top 10 speeches ranked by either the number of types or the number of tokens are listed in Table 2. There is a fair amount of overlap in the two ranking lists; in total only 13 speeches are in either of these rankings. Although the lengthiest speech (both in the number of words and the number of unique words) was given by Governor Mishkin, more than half of the top 10 speeches in either category were given by Governor Meyer. In addition, the top 10 speeches ranked by type-to-token ratio are shown in Table 3. Notably, there is no overlap with the speeches in the top 10 rankings by number of types and tokens in Table 2, revealing that the lengthiest speeches are not necessarily the most lexically diverse and emphasizing the importance of considering this additional metric.

[Insert Figure 3 here]

[Insert Table 2 here]

[Insert Table 3 here]

In addition to converting all words to lower case and removing punctuation and (numerical) symbols, the texts had to undergo a more extensive cleaning procedure, as is common in natural language processing, before proceeding to the formal analysis. Table OA1 of Online Appendix A gives an overview of the steps of the cleaning procedure along with the respective number of (unique) words in the texts after these steps. First, the inflected forms of the words are grouped together through lemmatization (e.g., 'bankers' is considered equivalent to 'banker').[11] Second, the texts were cleared of stopwords.[12] Third, words with a length of less than three characters were removed. Fourth, frequently co-occurring words (known as 'collocations'), were detected and grouped together via an underscore ('_') in order to be

[11] In contrast to stemming (another popular method to reduce inflected forms of words) lemmatization takes into account whether a word is a noun, adjective, adverb, or verb and brings the word back to a basic form that is still interpretable (i.e., that is still an actual word in the English language). As part of our interest is the actual words that comprise the topics, for the specific purpose of this paper, it is better for words to be reduced to a basic form that is still interpretable. If stemming were used, 'indeed' would be reduced to 'inde', which is somewhat difficult to interpret. Moreover, stemming comes with the risk of grouping too many words to the same basic form (e.g., 'marketing' and 'markets' both become 'market'), which is avoided when using lemmatization. With lemmatization, words can still be grouped by hand so that words with similar roots are treated collectively but the user has more control over how such grouping occurs.

[12] Stopwords are common words in the English language (e.g., 'and', 'the'). In this paper the stopwords list from the NLTK package version 3.2.1 in Python is used. In addition, words that are commonly used in describing numbers or statistics (i.e., 'number', 'percent', and the first 10 ordinal numbers) were added to the stopwords list. A list of all stopwords is included in Online Appendix A, Table OA2.

considered as one word in the text analysis.[13] Finally, the texts were cleared from words that occur in more than 95 per cent of the speeches (i.e., in more than 1056 speeches). A list of the words that were pruned in this final procedure is included in Online Appendix A, Table OA3.

3.2 Preliminary analysis

A self-constructed dictionary, which we will refer to as the 'Mandate Dictionary', is used to gain some preliminary insight into the use of words that might be related to the different

[13] When a collocation (i.e., a phrase) consists of two words it is referred to as a 'bigram'. For example, the words 'balance' and 'sheet' co-occur together frequently as the bigram 'balance sheet'. To detect these bigrams in the texts in our sample, the 'phrases' method in the Python gensim package is used. With this bigram detection method the two words that make up the bigram are merged together via an underscore into a single word. That is, 'balance' and 'sheet' are merged into 'balance_sheet' when the score of this phrase exceeds a certain threshold. The score of the phrase is computed by comparing the number of instances of the phrase itself to the number of instances of the words it consists of and this score is compared against the threshold value: (count(word *a* followed by word *b*) – minimum count)) * *N*/ (count(word *a*)*count(word *b*)) > threshold, where *N* is the number of unique words in the corpus (i.e., the number of unique words after removing punctuation, (numeric) symbols, lemmatization, and removing stopwords and words with a length below three characters. The parameter 'minimum count' is added to prevent too many phrases that consist of infrequent words from being created – that is, bigrams whose count does not exceed this minimum are ignored. In this paper a minimum count of 5 (the default value in the gensim package) and threshold value of 30 are used. The gensim package uses 10 as the default threshold value, but we have decided to use a more restrictive threshold of 30 in this paper. A lower threshold would allow for more bigram phrases and consequently increase the size of the vocabulary of the corpus, rendering the methods employed in this paper more computationally expensive. By computing the score of each possible word combination and comparing this against the set threshold, a final set of phrases is determined, which are merged together in the texts via an underscore. For more details on bigram detection see: https://radimrehurek.com/gensim/models/phrases.html.

mandates.[14] This small dictionary consists of three categories, intended to capture concepts corresponding to the dual mandate of the Federal Reserve: price stability, maximum employment, as well as the possible third one of interest, financial stability. The words in each of the dictionary categories are displayed in Table 4. We recognize that one of the risks of using a self-constructed dictionary is the possible inadvertent omission of certain relevant words, which is why we view it as merely an initial glance before we proceed with more advanced methods of text categorization, such as topic models.

The number of dictionary words per speech is positively related to the length of the speech, so we normalize by the total number of tokens in the speech in order to draw comparisons across speeches of different lengths. There has been an increase in the average percentage of financial stability and maximum employment words in speeches after the 2007-8 financial crisis, as shown in Figure 4.[15] This suggests there may have been a shift in the balance related to the Fed's mandate, with increased emphasis on the goals of maximum employment and controlling risks during/after the crisis, in contrast to an earlier emphasis on price stability. The timing of this shift coincides with the period of very low inflation that the US economy has experienced over the past decade.

[Insert Table 4 here]

[Insert Figure 4 here]

[14] Pre-existing dictionaries (e.g., the Loughran and McDonald (2011) dictionary for categorization of financial texts, or the Harvard IV-4 sentiment dictionary) have predefined classification categories (e.g., "superfluous", "litigious"), and therefore do not permit categorization of the speeches in a different manner, for example in terms of the Fed mandate.

[15] Scatterplots with the percentage of words per category over time are displayed in Online Appendix A Figure OA1.

As the results and associated inference may be sensitive to our choice of words in the constructed Mandate Dictionary, the speeches are re-evaluated using the financial stability words from the Peek et al. (2016) paper. The results of this robustness check are displayed in Figure 5. The pattern is similar to what was observed using our self-constructed dictionary: there appears to have been a post-crisis shift in the Federal Reserve's speeches, toward financial stability and maximum employment words, and away from price stability words.

[Insert Figure 5 here]

4. Methods

In order to extract the topics from the speeches we employ two different probabilistic models: the static Latent Dirichlet Allocation (LDA) model and its dynamic adaptation, Dynamic Topic Modelling (DTM).[16] One of the key assumptions of Latent Dirichlet Allocation (LDA) is that the order of the documents in the sample is irrelevant. This assumption is less palatable in the context of the particular subject of this paper, that is, an evaluation of whether the Federal Reserve's mandate implicitly has changed over time. For this paper, the order of the speeches is an important element to the determination of a changing focus in mandate and hence the DTM is more appropriate for estimating the evolution of topics over time within our collection of documents. Nevertheless, as LDA is the most commonly used topic model and has served as a source of inspiration for many other topic models, including the DTM (e.g., Srivastava and

[16] We use Python software (version 3.5.2), in combination with the natural language processing package Gensim (version 3.0.0) to obtain the results.

Sahami, 2009), we will first explain the LDA model in this methods section, and then explain the adaptations that were made to this model to get the DTM.

Section 4.1 introduces the terminology and notation. For both LDA and DTM the underlying assumption of how documents are created – the so-called *generative process* – is important to understanding the model. For this reason, both sections 4.2 and 4.3 lay out the steps of the generative process for LDA and DTM, respectively. Finally, section 4.4 explains how this generative process can be used to extract the topics and explore the ways the speeches exhibit them.

4.1 Terminology and notation

In the verbiage of natural language processing, the data to be used in our analysis are the words themselves. The dataset is the collection of documents to be analyzed, more formally referred to as *corpus*. The corpus consists of D documents (the 1173 speeches), each indexed by $d = 1, \dots, D$. The V number of unique words (types) in a corpus (that is, the union of all of the unique words in all the documents in the corpus) make up the overall vocabulary, and are each indexed by $v = 1, \dots, V$. After the extensive cleaning procedure described in the previous section, the size of the vocabulary is V=38,114.[17]

[17] The lemmatization procedure, the removal of stopwords and words with a length below 3, and pruning high frequency words all reduce the number of unique words, whereas the bigram creation procedure adds a lot of new unique "words" (i.e., the bigrams) to the sample. This is why this number differs from the number of unique words in the original dataset (24,116). The breakdown is presented in Table OA1 of Online Appendix A.

Every document in the corpus is a sequence of N_d words denoted by $\boldsymbol{w}_d = (w_{d,1}, w_{d,2}, \ldots, w_{d,N_d})$, where $w_{d,n}$ is the nth word of the dth document. The collection of documents in the corpus is formally defined by $\boldsymbol{w}_{1:D} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_D\}$.

All the documents in the corpus $\boldsymbol{w}_{1:D}$ share the same set of K topics, each indexed by $k = 1, \ldots, K$.[18] Every single word in the vocabulary occurs in each of the K topics with a different probability. More formally stated, $\beta_{k,v}$ is the probability with which the vth word in the vocabulary occurs in the kth topic. Hence, each of the K topics can be represented as β_k, a V-dimensional row vector of probabilities $\beta_{k,v}, v = 1, \ldots, V$ and can be interpreted as a discrete distribution over the V words in the vocabulary. This topic-word distribution β_k is also referred to simply as the topic. The words that occur with the highest probability in a topic determine what that topic is 'about'. All of the K topics can be concatenated into the $K \times V$ matrix β, where each row corresponds to one of the topics β_k.

Each document is assumed to be composed of a mixture of the K topics. For each topic β_k in document $\boldsymbol{w}_d$, the topic proportion is denoted by $\theta_{d,k}$. Analogous to the topic-word distributions, β_k, each of the d documents can be represented via a K-dimensional row vector of probabilities θ_d, $d = 1, \ldots, D$, that is, as a document-topic distribution over the K topics. These row vectors can be similarly vertically concatenated so that all of the D document-topic distributions of the corpus $\boldsymbol{w}_{1:D}$ are included in the $D \times K$ matrix θ, where each row

[18] For example, one would use K=2 if the only topics in the speeches were the two aspects of the Fed's dual mandate or K=3 to consider the possible emergence of a third aspect. More generally, it is assumed that the documents in the *corpus* cover a total of K topics, that no other topics are addressed in the documents, and that all the words in the documents occur in all of the K topics with a particular (nonzero) probability. Because the probability of a topic can be arbitrarily small, the larger K=3 encompasses the setting where a smaller number of topics is dominant.

corresponds to one of the document-topic distributions θ_d. More information about the topic-word and document-topic distributions is provided in section 4.2 for LDA and in section 4.3 for DTM.

4.2 The generative process of Latent Dirichlet Allocation (LDA)

LDA requires two important assumptions: (1) a topic is defined as a distribution over a fixed vocabulary of words, and (2) a corpus is associated with K topics and each document in the corpus exhibits a mixture of these topics.[19]

The main challenge of LDA is to infer the latent topic-word and document-topic distributions from the observed words. These two distributions (i.e., the topics and how each document exhibits them) are part of the hidden topic structure of the corpus. To understand how the hidden and observed variables interact it is necessary to first understand LDA's generative process. Specifically, it is an unobserved probabilistic process that is assumed to have generated the observed words. By reversing the generative process in the estimation, LDA is able to uncover the hidden topic structure of the observed words through posterior probabilistic inference.

In the generative process the V-dimensional vector of topic-word probabilities β_k and the K-dimensional vector of document-topic probabilities θ_d are represented by Dirichlet distributions. The Dirichlet distribution is commonly used in LDA because of its favorable properties in extracting the posterior distribution, which is explained in more detail in Online Appendix B.

[19] This latter assumption is what distinguishes LDA from the classical mixture models, which assume each document can only exhibit one topic.

The steps of the generative process for creating a corpus of words are first summarized in the following list; each step is then explained in more detail.

(1) For each topic: Draw a topic-word distribution over the V number words in the vocabulary: $\beta_k \sim \text{Dir}_V(\eta)$

(2) For each document $\boldsymbol{w}_d, d = 1 \ldots D$ in the corpus:

   (a) Draw a document-topic distribution over the K number of topics: $\theta_d \sim \text{Dir}(\alpha)$.

   (b) For each word position n, d in document $\boldsymbol{w}_d$ (where $n \in \{1 \ldots N_d\}$ and $d \in \{1 \ldots D\}$):

      (i) Choose a topic from the document-topic distribution θ_d of step (2a) to assign to the word $w_{d,n}$. This is the topic-assignment: $z_{d,n} \sim \text{Multinomial}(\theta_d)$.

      (ii) Choose a word from the topic-word distribution corresponding to the topic-assignment of step (2bi): $w_{d,n} \sim \text{Multinomial}(\beta_{z_{d,n}})$.

The first step is to draw (determine) the Dirichlet topic-word distribution:

$$p(\beta_k|\eta) = \frac{\Gamma(\sum_{v=1}^{V} \eta_v)}{\prod_{v=1}^{V} \Gamma(\eta_v)} \prod_{v=1}^{V} \beta_{k,v}^{\eta_v - 1}$$

where η_v is the weight of the vth word in the vocabulary in topic β_k, and is often referred to as the *hyperparameter* of the distribution. The expression $\frac{\Gamma(\sum_{v=1}^{V} \eta_v)}{\prod_{v=1}^{V} \Gamma(\eta_v)}$ is the normalizing constant of the Dirichlet distribution (as it is used to guarantee the distribution has a total probability of one),

and Γ denotes the Gamma function.[20] Step 1 is repeated K times, once for each of the K topics in the corpus. The next step (2a) is to draw a document-topic distribution for document $\boldsymbol{w}_d$:

$$p(\theta_d|\alpha) = \frac{\Gamma(\sum_{k=1}^{K} \alpha_k)}{\prod_{k=1}^{K} \Gamma(\alpha_k)} \prod_{k=1}^{K} \theta_{d,k}^{\alpha_k - 1}$$

where α_k is the weight of the kth topic in the document-topic distribution θ_d. The two Dirichlet distributions, with hyperparameters η and α respectively, are labelled the *prior distributions*, as they capture the initial beliefs about the topic-word and document-topic probabilities (Steyvers and Griffiths, 2006).

Once the topic distribution has been determined, the word generation process can start. Let $z_{d,n}$ be the topic that is assigned to the position of the nth word in the dth document.[21] The probability of topic-assignment $z_{d,n}$ given the document-topic distribution (the multinomial in step (2bi)) is: $p(z_{d,n}|\theta_d) = \theta_{d,z_{d,n}}$. The topic-word distribution corresponding to topic-assignment $z_{d,n}$ is $\beta_{z_{d,n}}$. From that topic-word distribution a word $w_{d,n}$ is chosen with probability $p(w_{d,n}|z_{d,n}, \beta) = \beta_{z_{d,n},w_{d,n}}$. Steps (2bi) and (2bii) are repeated N_d number of times

[20] More precisely, the normalizing constant is the inverse of the multivariate Beta function: $B(\eta) = \frac{\prod_{v=1}^{V} \Gamma(\eta_v)}{\Gamma(\sum_{v=1}^{V} \eta_v)}$. The Beta function is expressed in terms of the Gamma function Γ,and the definition of the Gamma function is: $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$. For positive integers, this Gamma function is an extension of the factorial function and can be written as: $\Gamma(z) = (z-1)!$.

[21] More specifically, topic-assignment z is a K-dimensional unit vector with components z^k, $k = 1, \dots, K$, such that $z^k = 1$ and $z^i = 0$ for all $i \neq k$. The probability that a topic is assigned to a particular word depends on the document's distribution over the K topics. The set of topic assignments in document $\boldsymbol{w}_d$ is denoted by $\boldsymbol{z}_d = (z_{d,1}, z_{d,2}, \dots, z_{d,N_d})$ and the collection of topic assignments in the corpus is $\boldsymbol{z}_{1:D} = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_D\}$.

until each word position of the document has been filled. This concludes the creation of a document. Steps 2a and b are then repeated D number of times until the entire corpus is generated. A summary of the terminology discussed in sections 4.1 and 4.2 is included in Table 5.

[Insert Table 5 here]

The notation and dependencies between the variables of the generative process are displayed in the graphical model in Figure 6 (Blei et al., 2003; Blei, 2011).

[Insert Figure 6 here]

Each node depicts a random variable: shaded means the variable is observed and unshaded means it's latent. The rectangular boxes ('plate notation') in the graphical model are used to indicate replication of particular nodes (Srivastava and Sahami, 2009). This graphical model shows that the topic assignment $z_{d,n}$ of word $w_{d,n}$ in document $\boldsymbol{w}_d$ depends on the document's topic proportions θ_d. Similarly, the observed word $w_{d,n}$ depends on the topic assignment $z_{d,n}$ and *all* of the K number of topics β (Blei et al. 2003; Blei, 2011). The figure also reveals that the variables α, and η are *corpus level variables* (global variables), assumed to be sampled only once during the corpus generating process. The variable β_k is also a global variable, sampled not once but K number of times, i.e., once for each of the topics in the corpus. The variables θ_d are *document-level* (local variables) and are sampled once for each of the D documents in the corpus. Finally, $z_{d,n}$ and $w_{d,n}$ are sampled once for each of the N_d words per document, making them *word-level variables* (Blei et al. 2003).

4.3 The generative process of the Dynamic Topic Model (DTM)

As noted earlier, for the purposes of evaluating the evolution of topics over time, the Dynamic Topic Model (DTM), a dynamic version of the static LDA model, is more appropriate. By adapting state space models to categorical data, Blei et al. (2006) were able to incorporate the time dynamics of a corpus' underlying topics.

A corpus is split into T time slices, and contrary to the notation employed in LDA, D is used to indicate the number of documents in a time slice (not the entire corpus). Hence, the notation introduced in the previous section (see Table 5) is used for the DTM model as well, but a subscript t is added whenever it concerns a time slice as opposed to the entire corpus. In this paper, yearly time slices (i.e., t is year) are used, which means that the topics associated with a particular year are assumed to have evolved from the topics the year before.

As the Dirichlet distribution is not suitable for sequential modelling (Blei et al., 2006), the natural parameters of each topic β_k at a given point in time, $\beta_{k,t}$, are chained together via a state space model that evolves with Gaussian noise (Blei et al., 2006; Srivastava and Sahami, 2009).[22] That is, the time-varying topic-word distribution $\beta_{k,t}$ is a multivariate Gaussian random variable for the kth topic in time slice t, with mean $\beta_{k,t-1}$ and chain variance parameter σ^2 that determines how fast the topic evolves over time. In contrast to LDA, $\beta_{k,t}$ is not a distribution over words, but a real-valued V-dimensional vector of parameters. When words are drawn from this vector of natural parameters (step (2bii)) they are first mapped back to a set of vectors that sum to 1 through the function f. That is, the representation of a topic as $f(\beta_{k,t})$ ensures that

[22] The Gaussian density is chosen because its symmetry properties enable the use of standard forward-backward calculations for linear state space models (Blei et al., 2006).

each topic is a distribution over the words in the dictionary, and that all the word probabilities sum to one (as was the case with LDA as well). More specifically, a word $w_{d,n}$ is chosen with probability:

$$p(w_{d,n}|z_{d,n},\beta) = f(\beta_{z_{d,n},t}) = \frac{\exp(\beta_{z_{d,n},w_{d.n},t})}{\sum_v \exp(\beta_{z_{d,n},v,t})}.$$

The generative process for the D number of documents in the sequential corpus in time slice t is:

(1) Draw topic-word distributions $\beta_{k,t}|\beta_{k,t-1} \sim \mathcal{N}(\beta_{k,t-1}, \sigma^2 I)$, where σ^2 denotes the chain variance.

(2) For each document $\boldsymbol{w}_d, d = 1 \ldots D$ in time slice t: proceed as in step (2) of the LDA model, with the distribution in step (2bii) now being $w_{d,n} \sim$ Multinomial$(f(\beta_{z_{d,n},t}))$.

The dependencies between the variables are displayed in the graphical model in Figure 7. Both the generative process and the graphical model show that each time slice corresponds to a separate LDA model, of which the topic-word distributions are allowed to evolve from time slice to time slice.

[Insert Figure 7]

4.4 Extracting and exploring the topic structure

Assuming the hidden topic structure of a corpus is represented by the topics (β), the document-topic distributions (θ) and per-word topic assignments ($\boldsymbol{z}_{1:D}$) can be estimated using Bayesian updating. The posterior distribution is essentially the update given to the prior beliefs (i.e., distributions) after observing the corpus. More specifically, given an observed corpus (and

prior topic-word and document-topic distributions β and θ, with parameters η and α, respectively), a posterior distribution over the hidden topic structure variables can be obtained: $p(\theta, \beta, \boldsymbol{z}_{1:D}|\boldsymbol{w}_{1:D})$. This posterior distribution can be thought of as the 'reversal' of the generative process that created the corpus (Srivastava and Sahami, 2009). As the true posterior distribution is intractable to compute, it needs to be approximated (Blei, 2003). In this paper, mean field variational inference is used as the approximation method for both LDA and DTM, as that is the preferred approach to approximate the posterior in the DTM (Blei, 2003; Blei et al. 2006). In general, variational inference is especially useful when analyzing large datasets, which is often the case in text analysis (Grimmer, 2011). The mean field variational inference technique provides a variational distribution that is closest to the posterior distribution. Before the variational distribution can be estimated, a number of items first need to be specified. These include the prior distributions (i.e., the two Dirichlet distributions in LDA with prior weights η_v and α_k, respectively) and the number of topics K.[23,24] As this paper investigates the presence of

[23] In our approximation of the posterior distribution, α and η are treated as constants, thereby employing a symmetric Dirichlet prior which assumes equal prior probabilities for all the components in the vectors α and η. More specifically, α is set to $50/K$ and η to 0.01, as those values have worked well with many different text collections (Steyvers and Griffiths, 2006). For comparability, the same values of α and η are used for LDA and the DTM. In the Variational Inference estimation of the DTM, the LDA model is used as initialization of the first time slice, which is why the prior choice of not just α, but also η, is relevant for the DTM.

[24] In the DTM a value additionally needs to be set for the chain variance parameter as well (σ^2). The higher this chain variance is, the more words within the topics are allowed to vary over time. In this paper the default value of 0.005 is maintained, as recommended by Blei et al. (2006). As a robustness check, an increase by one order of magnitude was explored but that did not result in large changes to the results.

three mandates, the number of topics is set to $K = 3$.[25] For more details on mean field variational inference, see Online Appendix B.

The quantities required to study the topic structure of a collection of documents can be readily obtained from the variational distribution. The topics of a corpus are explored by inspecting the topic-word probabilities of each word per topic (also referred to as the term probabilities): $\hat{\beta}_{k,v} = E\left[\beta_{k,v} \middle| \boldsymbol{w}_{1:D}\right]$. A simple way to rank topics is to order these probabilities from highest to lowest, thereby revealing which words are most likely to occur in that topic. As LDA (and DTM) allows terms to occur in multiple topics, another possibility is to order the terms in a topic by term-score:

$$\text{term-score} = \hat{\beta}_{k,v} \log\left(\frac{\hat{\beta}_{k,v}}{\left(\prod_{k=1}^{K} \hat{\beta}_{k,v}\right)^{1/K}}\right)$$

The term score down-weights terms that have a high probability under all the topics, thereby assigning a higher weight to the discriminative words in a topic. Ordering the words by term-score reveals the words that are most distinctive for a particular topic.

In addition to inspecting the topics, the corpus can be explored by inspecting the posterior document-topic proportions of each document: $\hat{\theta}_{d,k} = E[\theta_{d,k} | \boldsymbol{w}_{1:D}]$. Ordering these topic-

[25] Another way to determine the number of topics would be to use the coherence score, which represents the human interpretability of a topic (Mimno, Wallach, Talley, Leenders, McCallum, 2011). The higher the score, the more interpretable the topics are for people. When estimating an LDA model on the entire corpus allowing the number of topics to vary from three to 10, the highest coherence scores were obtained for the LDA models with 2, 3 and 4 topics. As the scores did not differ much from one another (-0.200, -0.267 and -0.317 for 2, 3 and 4, respectively) and as this paper focusses on examining the presence of three versus two mandates, the number of topics was set to three.

proportions of each document from highest to lowest reveals the main topics of a specific document.

On a final note, because lemmatization is used in the data cleaning process, words that are related are at times not reduced to one base form. For example, 'inflationary' is not reduced to 'inflation' and 'joblessness' is not reduced to 'jobless', as lemmatization takes into account whether words are nouns, adjectives, adverbs, or verbs. However, for our purposes in a number of instances, it is more informative to aggregate the per-topic term probabilities of related words (for example, if the words 'economic' and 'economy' both end up in the top 10 words of a topic, it is more interesting to aggregate these into a single probability and see what other words rise to the top 10 as a result). Related words are therefore grouped and their respective per-topic term probabilities are aggregated. Online Appendix A lists an overview of the constructed groups in Table OA4. In the word rankings of the topics only the first word of each of these word groups is shown in the results section (e.g., the group 'inflation/inflationary' is displayed as 'inflation').

5. Results and Discussion

As the DTM is better-suited to gauge the extent to which the topics change over time, we will first discuss the DTM results in section 5.1, and then compare those to the results obtained when estimating the LDA model for each year separately in section 5.2.

5.1 DTM

Tables 6 to 8 display the top 10 words for each of the 3 topics from 1997 until 2016, ranked in two ways. Panel (a) ranks the words by their occurrence probabilities and panel (b) by

their term scores. Because the latter facilitate topic discrimination, the discussion in this section focuses on the term-score results.

5.1.1 Topic 1: Economic and Business Developments

Table 6 shows the words associated with topic one – these might be classified as relating to (*global) economic and business developments*. In the first half of the sample period, this topic was defined by words concerning business and technology (e.g., 'technology', 'business', 'investment'), economic, labor and productivity growth (e.g. 'economy', 'growth', 'increase', 'labor', 'work', 'productivity', 'productivity growth', and 'production') and household spending (e.g., 'household', 'spending'). These words reflect technological changes and accompanying changes in the workforce during that time, concern over Y2K, and the bursting of the tech bubble that took place from 1997 to 2001. During the early years of the global financial crisis (2007-2009) the topic changed and began to emphasize phrases related to the credit crisis (e.g., 'housing', 'home', 'mortgage', 'house', 'house_price', 'unemployment'), with the word 'recovery' first appearing in the top 10 list in 2010. Finally, from 2012 the topic shifted again more towards the financial needs of households, with an increased emphasis on employment opportunities and housing (e.g., 'work', 'job', 'household', 'income', 'spending', 'housing', 'unemployment', 'labor' and 'labor_force'). In those years the words 'inflation', 'economy', 'growth', and 'recovery' started to increase in probability and/or term score as well, signaling economic recovery and the possible commencement of the tightening cycle. Interestingly, throughout the entire sample period words concerning employment (e.g., 'productivity', 'work', 'labor', 'job') occurred in this topic with a high probability/term score. Before the crisis the speeches covering this topic honed in on the productivity growth due to new technologies,

whereas in the years after the crisis more emphasis was put on the unemployment levels and job opportunities. In addition, the top 10 word lists reveal a change in the Fed's approach to the employment mandate; up until 2006 the focus was on productivity growth, whereas after 2010 this shifted toward tackling unemployment.

The evolution of the term scores of words that occurred in the top five at least once from 1997 to 2016 is displayed in Figure 8.[26] It is evident that there is a lot of variation in the top words for this topic. In the years of the financial crisis the topic was defined more by words 'home', 'house_price' and 'housing', whereas the words 'productivity', 'productivity_growth', and 'technology' had been more important in earlier years. Moreover, before the crisis, the topic focused more on (oil) prices, economic output, the current account deficit and investments, whereas during the crisis 'spending' was important, as well as words related to housing. In later years the focus was on economic growth and employment opportunities ('job', 'labor', 'unemployment', 'recovery', 'growth'). This reveals that the Fed's communication on this topic has been largely consistent with its actions: during the financial crisis the Fed initially focused more on consumers' financial and housing situation, whereas in recent years it has shifted its focus to employment and income (i.e., its employment mandate).

[Insert Table 6a here]

[Insert Table 6b here]

[Insert Figure 8 here]

[26] An earlier version of this paper also contained analogous figures of words ranked solely by their occurrence probabilities. These are included in Online Appendix C, Figure OC1.

5.1.2 Topic 2: Monetary Policy

Tables 7a and 7b show the top 10 words per year associated with topic two, determined by their probability and term score rankings, respectively. This topic can be labelled as *monetary policy*. Figure 9 displays the evolution of the term scores of the top five words over time, and these show that there is less variation over time in the top five words for this topic than for topic one. The graphs show sudden peaks in the probability and term score of the word 'inflation' in 2000 (when the Fed reached the peak of its tightening cycle before aggressively lowering the federal funds rate below 2 percent in the following year), in 2003 (as the Fed was at historically low levels of the policy rate and the market began to fear the Fed would not tighten soon enough), from 2005 to 2008 (during the Fed's tightening cycle), and a drift upward in 2015 and 2016 (as the Fed ended its prolonged episode at the zero lower bound). In the second half of the sample, especially from 2012 until 2016, a lot of attention has been paid to the interest rates and the federal funds rate. Specifically, there has been a decrease in the probability and term score of the words 'price', 'target' and 'inflation', and an increase in the words 'fund', 'rate', 'interest_rate', 'committee', and 'FOMC'.[27] These shifts reveal that before the start of the financial crisis in 2007, the focus was on ensuring price stability but starting 2012 the focus shifted towards an emphasis on moderate longer term interest rates. Increased importance of the word 'financial' in Table 7a, coinciding with an increase in the probability of words related to the Federal Reserve ('Federal_Reserve', 'Fed', 'FOMC') in Table 7b, reveals how the Federal Reserve's monetary policy decisions and the (consequences for the) financial system have become more intertwined over the years. All in all, the changes in this second topic over the

[27] In the cleaning process all words were lower-cased, but in the display of the results proper nouns are expressed with the correct capitalization (e.g., Federal Reserve, FOMC, etc.).

years show changes in the Fed's approach to its price stability mandate. Interestingly, in 2004, 2005 and 2006, and again in 2008, 2009 and 2010 (i.e., the years leading up to the crisis and in the early years of the financial crisis) there was a slight decrease in the term score of the word 'monetary_policy' in Figure 9, but in the years from 2012 to 2016 the probability returned to the same level as before the crisis.

[Insert Table 7a here]

[Insert Table 7b here]

[Insert Figure 9 here]

5.1.3 Topic 3: Risks, Regulation, and Supervision of the Banking System

Tables 8a and b show the words associated with topic three. These can be labelled as words associated with *risks, regulations and supervision of the banking system*. The evolution of the term scores of words that occurred (at least once) in the top five over the years is displayed in Figure 10. The top 10 lists and the figure show that there are clear changes between the period before the financial crisis and thereafter. In the years leading up to and at the start of the financial crisis, the words 'risk', 'Basel', 'institution', 'management', 'credit', 'small_business', 'loan' and 'banking' were important, coinciding with the Basel agreement and the legislative and regulatory changes in the Community Reinvestment Act in 1999 and 2005, but the probabilities and term scores of those words have decreased in recent years. In contrast, the words 'firm', 'financial', 'systemic', and 'system' have become more important, perhaps reflecting the policymakers' shift in focus from microprudential to macroprudential supervision. In the first half of the sample the topic focused more on supervision of an individual bank, whereas in the second half the language shifted towards (capital) regulations and crisis prevention, as is shown

by the decrease in the word 'supervision' and increase in the words 'systemic', 'regulation', 'capital', and 'requirement'. More specifically, Figure 10 reveals that the words 'requirement', and 'regulation' have gained importance in recent years. In addition, Tables 8a and b show that the words 'lender', 'loan', 'credit' and 'mortgage' and 'liquidity' are prevalent in the rankings in the second half of the sample. Interestingly, as shown in section 3.1, many of these words have been connected to the presence of a financial stability mandate (Peek et al., 2016). But it is not just the presence of these words that reveals a changed approach of the Fed to the financial system. Over the entire sample period the word 'bank' is at the top of the lists (highest term score every year, and top 3 of the topic-word probabilities) but starting in 2009 the word 'firm' suddenly appears in all the yearly lists as well. Moreover, the word 'system' disappeared from the probability-ranked lists after 2002 but reappears in 2010 and all the years thereafter. The increased importance of the words 'firm', 'systemic' and 'system', along with the rise in the importance of the word 'community' and the discussion on capital requirements and liquidity regulations, reveals the wider approach the Federal Reserve is taking in ensuring the stability of the financial system post-crisis. Even though the topic over the whole sample cannot be labeled 'financial stability' (as that does not describe the words in the top 10 lists in the first half of the sample), the words that define this topic in recent years reveal that the Fed's discussion of regulations and supervision in the banking system has shifted to a financial stability focus.

[Insert Table 8a here]

[Insert Table 8b here]

[Insert Figure 10 here]

5.1.4 Word Clouds of the Top 100 Words in Each Topic

Figure 11 displays the word clouds of the top 100 unique words of each of the three topics before, during, and after the crisis. Specifically, the word clouds are constructed from the top 100 words in each topic after reordering the rankings of the words such that a word only occurs in the topic in which it had the highest probability or term score. That is, if a word is initially in the top 100 of multiple topics, it is omitted from the top 100 of the topic(s) in which it had a lower probability/term score and the next subsequent word (that does not have a higher probability/term score in another topic) is included in the top 100.[28] The color of the words indicates the topic it belongs to and the size is proportional to the average probability (left column) or average term score (right column) of that word over each subperiod. That is, if a word has twice the probability (or term score) of another word, it is twice the size in the word cloud. Hence, these figures provide a different visual representation of the results displayed in Tables 6 to 8 and Figures 8 to 10.

The word clouds based on word probabilities (left column) show that before the crisis (top cloud) the speeches focused on risk management, economic growth, inflation, prices, banking supervision, and interest rates. During the crisis (middle cloud), there was still a large emphasis on inflation policy, but in addition to that the speeches put emphasis on the state of the economy and financial conditions, credit, risk, mortgages, housing, and consumer loans. In the years after the crisis (bottom cloud), regulation became a more dominant focus, as the emphasis

[28] Figure OD1 in Online Appendix D shows the word clouds with duplicate words, where the topic number is appended to the word to indicate whether, for example, the word 'bank' corresponds to topic 1, 2 or 3. These word clouds confirm the findings in Figure 11.

of the speeches shifted towards financial regulations, (financial) policy, capital risk, unemployment, and the federal funds rate.

The right column of Figure 11 displays the word clouds based on the term scores. What becomes immediately clear is that the relative term score differences between words are much larger than the relative word probabilities. This is because term scores put more weight on the more distinctive words in texts. The clouds show clearly that the word 'inflation' and the word 'monetary_policy' are discriminative words for the monetary policy topic. Moreover, the word clouds show that before (and somewhat during) the crisis the word 'supervision' was highly prevalent in the risk and regulations topic, whereas after the crisis the word 'regulation' grew in importance, indicating a shift from a relatively passive to a more proactive approach to managing the risks in the banking system. Finally, the term-score word clouds (right column) confirm previous findings that from 2007 until 2009 the speeches were more focused on the mortgage crisis itself ('housing', 'house_price', 'mortgage', 'loan'), while after the crisis more emphasis was put on the need to create jobs and decrease unemployment.

[Insert Figure 11 here]

5.1.5 Document-topic Proportions

Figures 12 and 13 show the document-topic proportions estimates. The document-topic proportions provide a way to gauge how the importance of these topics within the speeches has evolved over time. Figure 12 displays the average document-topic proportion across all speeches in a given year. Topic one, corresponding to economic and business developments, appears on average to have a slight cyclical movement over the years, with noticeable dips in the years of the financial crisis and in 2013. In the last years, its importance has started to increase again,

coinciding with an increase in the probability of words associated with the employment mandate. While the coverage of the first topic declined somewhat in the second half of the sample compared to the first half, the second and third topics became more prevalent in speeches. The monetary policy topic especially has received extensive coverage in speeches in 2013, 2014, 2015 and 2016, coinciding with the increased importance of words associated with the FOMC and the federal funds rate. The year 2013 includes the period when Treasury yields rose sharply, in response to the Fed's `taper tantrum'. The increased importance of the risks, regulations, and supervision topic in the years of and following the financial crisis, together with the increased probability and score of words related to mitigating systemic risk, capital regulations and liquidity requirements, serves as a sign of an emerging third mandate. In addition, Figure 12 reveals that not only was this third topic important from 2008 onwards, but it briefly enjoyed the focus of a number of speeches back in 1997, 1998 and 1999 as the original Basel capital regulations went into effect. In those years the language of the topic focused on risk, loans, and supervision. Not surprisingly in 2006, in the year leading up to the November 2007 passage of the US version of the Basel II capital regulations, the topic also received a lot of attention in speeches. In 2005 and 2006 the term score of the topic three words 'Basel' and 'risk' noticeably increased from 0.009 and 0.016 in 2004, to 0.017 and 0.019 in 2006, respectively (as seen in the rankings of Tables 8a and b and in Figure 10).

[Insert Figure 12 here]

Overall, Figure 12 shows that the dominant topic in most of the speeches in each year was the one concerning risks, regulations, and supervision in the banking system, topic 3. On average, almost 60 percent of every speech in 1997 and 1998 was devoted to this topic,

coinciding with the legislative developments surrounding the repeal of the Glass-Steagall Act and the passage of the Community Reinvestment Act. In those years – and in 2000, 2001, and 2002 – words like 'loan', 'community', 'small_business' and 'financial', 'institution', 'banking' and 'system' were important. In contrast, in 2004 and 2005 it was more likely that a speech would cover economic developments. Specifically, the speeches in those years focused on rising oil prices, labor and productivity growth, the rise in investments and the current account deficit. In 2006, the dominant topic in most speeches was again the third topic (roughly 50 percent), coinciding with the release of the Fourth Quantitative Impact Study assessing banks' preparedness for risk-based capital regulations (`Basel II') and associated negotiations leading up to the 2007 passage of these regulations. A similar high percentage was noticeable for the third topic in 2008, 2009, and 2010, the years of the credit crisis, when banks, mortgages, loans and new regulations were the subject of discussion. In 2013, 2014, and 2015 the third topic concerned the liquidity regulations and capital requirements that were aimed at protecting the financial situation of communities and preventing another crisis. In those final years of our sample the speeches were more likely to discuss either the monetary policy or the risk, regulations and supervision topic, a pattern we have seen in the previous figures as well. Specifically, those speeches focused on the federal funds rate, interest rates, and economic growth, as the Fed sought to begin its tightening cycle. During this time, there also was an increased emphasis on the financial stability, as indicated by the higher word probabilities/term scores for 'firm', 'systemic' and 'system' in Tables 8a and 8b.

One challenge with reporting proportions across speeches is that if speeches are highly unidimensional (i.e., have low document-topic proportions for all except one topic), fluctuations in Figure 12 could merely reflect the proportion of speeches that do not cover a given topic

rather than truly capturing fluctuating emphasis within speeches. Figure OC2 in Online Appendix C takes this into account by dropping (from the calculation of each topic's average) those speeches that have very low document-topic proportions (i.e., proportions below 0.05) for that topic and shows that the results displayed in Figure 12 are robust to this concern. In addition, Figure OC3 in Online Appendix C displays the percentage of speeches per year devoted to a given topic, where each speech is classified into only one of the three topics, according to which topic obtained the maximum document-topic proportion for that speech. The pattern under this latter classification is more pronounced but similar to the pattern observed before in Figure 12 (and OC2).

[Insert Figure 13 here]

However, as speeches usually do have one 'main' topic, the highest document-topic proportions are investigated in more detail in Figure 13, showing the percentage of speeches (over all the speeches in a given year) dedicated to a single topic, where topic-dedication is defined by a speech where one of the three topics had a document-topic proportion of over 0.95. Hence, it displays the percentage of speeches by year where just one of the three topics was covered extensively. The figure shows a large proportion of single-focus speeches associated with the banking risks, regulations and supervision topic (topic 3). In 1997, 2006, 2014 and 2015 approximately 30 percent of the speeches focused solely on that topic, coinciding with important changes in financial regulations in those years, as described above. In the first half of the sample period, there were more single-focus speeches on economic and business developments, which discussed disruptive technology changes, the current account deficit, rising (oil) prices, or the increase in investments. The years 1999, 2000, 2001, 2004 and 2005 are the

only years in which approximately 10 per cent of the speeches was solely focused on monetary policy, around the same time that the word 'expectation' turned up in the top 10 wordlists. The appearance of the word 'expectation' in the top 10 wordlists coincides with the start of the Fed's tightening period and a large spike in the proportion of times that the word 'expectation' is preceded by the word 'inflation' (see Figure 14). The spike in the 'inflation expectations' proportion also coincides with an increase in the Fed's focus on managing inflation expectations.[29] From the beginning of the sample through 2002, the percentage varies between one and eight percent, roughly tracking both the inflation rate as measured by the Consumer Price Index (year-on-year change) and the federal funds rate, reaching its lowest level in 2002. A double-digit proportion is first observed in 2003 when the Fed halted its easing cycle and remains elevated during the Fed's tightening period. The emphasis on inflation expectations increased even more in the run-up to the financial crisis; in 2007 and 2008 'expectation' was preceded by the word 'inflation' over 26 percent of the time, even as the Fed aggressively drove the policy rate to the zero lower bound. The percentage decreased as both inflation and the policy rate remained low, with a particularly large drop in 2011 but in 2015 and 2016 rose dramatically to double-digit levels as the Fed emerged from the zero lower bound and began its current tightening cycle.

[Insert Figure 14 here]

[29] See, e.g., the transcripts from the March 16, 2004, May 4, 2004, and June 29-30, 2014 FOMC meetings, Federal Open Market Committee (2014).

5.2 LDA

The LDA model is estimated for each year separately. Table 9 displays the interpretation given to the three topics for each year from 1997 until 2016. These interpretations are based on the word probability and term score rankings of the top 10 words per topic. While this section discusses inference from these rankings, in the interest of space the actual lists are given in Tables A1 to A20 in the Appendix. Table 9 shows that in almost all years one of the topics concerned economic and business developments, a second one concerned monetary policy, and a third concerned the risks, regulations and supervision in the banking system.[30]

[Insert Table 9 here]

In the first half of the sample the economic and business developments that were discussed in the speeches varied widely: community development lending services for

[30] It is important to note that this does not constitute the order of the topics in which they appear initially. In some years the first topic concerned monetary policy, whereas in other years it could be the second or third. As the order in which the the topics initially appear is random, tables D1 to D20 were reshuffled to ensure that topics with a similar interpretation would have the same topic number. Hence, the order in which the topic interpretations in Table 9 appear is a result of this reshuffling process, and the order is the same as the (order of the) topic interpretations of the DTM model. That is, the first topic is economic and business developments, the second concerns monetary policy, and the third one concerns risks, regulations, and supervision of the banking system. In years where one of these three main interpretations could not be given to one of the topics, the other topic interpretation would come in place of the missing one, as visible in Table 9. For example, in 2003 none of the topics could be interpreted as covering 'economic and business developments', but instead there was one covering 'internal control, corporate governance and compliance'. Hence, that topic is listed as the first topic that year.

households and small businesses (1997, 1998, 2002, and 2006), financial technology and new payment systems (1997, 1998, 1999, 2000, 2001), the current account deficit, rising energy prices, and productivity growth in the labor market (2000, 2001, 2004, and 2005). During the crisis years the topic that discussed economic and business developments had a high probability/term score for words concerning subprime mortgages, foreclosures and credit risks. From 2009 until 2015 this shifted more toward words related to community credit needs, such as loan programs for families and small businesses, jobs, and household income. In 2010 this economic and business developments topic discussed consumer finances (e.g., 'credit_card', 'credit', 'consumer', 'lender', 'consumer'), likely in response to the passage of the Dodd-Frank Act. Finally, it is interesting to note that in the second half of the sample words related to employment gained importance.

In most years monetary policy and inflation (targeting) was discussed in relation to the economic outlook. Before the crisis the word 'growth' often had a high word probability/term score, while during and after the crisis the word 'recovery' was more likely to occur in relation to the word 'economy'. In contrast to the DTM results, with LDA the top words of the monetary policy topic are not solely comprised of words related to the price stability mandate, but also of words related to the employment mandate. Monetary policy also often was discussed jointly with the unemployment rate and the state of the job market, even more so in the second half of the sample period. In the first half of the sample period the monetary policy topic was defined by words related to rising (oil and asset) prices and government policies such as social security (2000 and 2001), reflecting inflationary concerns that would have driven decisions concerning the cessation of the easing period (that began in mid-2000 and reached its floor in mid-2003) and the eventual commencement of the tightening cycle in June 2004 that continued until reaching a

plateau in June 2006. There were three years – 2001, 2006 and 2009 – in which the phrase 'monetary_policy' did not occur in the top-10 topic word lists, but where words like 'money', 'currency', 'rate', and 'interest_rate' reveal that speeches did refer to monetary policy issues. Finally, in the second half of the sample period the monetary policy topic was defined more by words related to the FOMC and the federal funds rate (2011, 2012, 2013, 2014, 2015 and 2016), and words related to the Fed's purchase of longer-term securities (2011, 2012, 2013, 2014).

Finally, the word lists in Tables A1-A20 show that the risks, regulations and supervision in the banking system topic focused on risk management (1997, 2000, 2005, 2006 and 2008) and the Basel II capital regulations (2001, 2003, 2005, 2006 and 2007) in the first half of the sample period. Words related to the (housing) crisis gained importance in the banking regulations and supervision topic in the years 2007, 2008, and 2009. Thereafter, the focus of this topic shifted more towards capital requirements, liquidity regulations, and stress tests. This pattern is very similar to the DTM results.

Even though in most years these three main interpretations (i.e., economic and business developments, monetary policy, and risks, regulations and supervision in the banking system) could be given to each of the three topics, there were some years where the interpretation of one of the topics differed from these three. More specifically, in both 2003 and 2004 one of the three topics covered internal control, corporate governance, and compliance within financial organizations and businesses; these years corresponded to the discovery of the Freddie Mac and Fannie Mae accounting fraud scandals, respectively, and followed the 2002 passage of the Sarbanes-Oxley Act in the aftermath of the Enron scandal in 2001. This internal control and compliance topic can be viewed as a combination of two of the three main topic interpretations (economic and business developments, and risks, regulations and supervision), as it combines

both developments in business as well as risk management. In both 2012 and 2014 one of the three topics was devoted to financial stability. More specifically, financial stability was discussed in relation to firms and financial institutions (2012), financial needs of families and communities (2012, 2014), and monetary policy (2014). This 'financial stability' topic also can be viewed as a combination of two of the main topic interpretations (the monetary policy and risks, regulations and supervision topics), as it encompasses both monetary policy measures, as well as systemic risks in the banking system.

Examining the top 10 words themselves further yields a number of interesting patterns. The following words are among the top 10 in at least one topic for all the years in our sample: 'bank', 'economy', 'financial', 'rate', 'risk' (for the top 10 lists based on word probabilities), and 'inflation' (for the top 10 lists based on term scores). For those words that are not in at least one topic in all years, there are additional interesting patterns associated with the first appearance of a number of words. For example:

- "CRA" (the Community Reinvestment Act), which appeared in 1997, 1998, 1999 coinciding with proposal of the Gramm-Leach-Bliley Act (GLBA), which passed in 1999. One of the requirements of the GLBA was that the Fed would conduct a study of CRA, to be delivered to Congress in early 2000 and hence it is not surprising that CRA was the focus of a number of speeches. The abbreviation appeared again in 2009, along with the phrase 'big_fail'[31], highlighting how companies had become too big to fail due to the GLBA.

[31] This phrase originally appeared as 'big to fail' in the texts, but because the stopwords (including the word 'to') were removed before the bigrams were created, the phrase became 'big_fail' in the bigram creation procedure.

- "Basel" first appears in 2001 and rises to the top of the word list by term score in 2005; these are the years after the European Capital Requirements Directive (CRD) and Basel II, respectively, went into effect in Europe, with the US expected to follow suit. In April 2005, the US regulatory Agencies[32] announced a delay to Basel II implementation over concerns that minimum regulatory capital would fall precipitously under the new proposed regulations. That was followed by months of highly contested negotiations among the agencies, often aired quite publicly.
- 'Derivative' and 'option' appear for the first time in 2002 and 'mortgage' appears for the first time in 1998, quite a few years before the start of the financial crisis.
- 'Credit' appears in more than half of the years in our sample (1999, 2002-2005, 2007-2014) and long before the start of the financial crisis (a.k.a. the credit crisis) in 2007.
- 'Stability' occurs for the first time in 2003, in relation to the economic growth of that time and the rise in inflation. Thereafter the word occurs in 2012 and 2014, in the context of financial stability.
- "Recession" already appeared in the top 10 term scores of the monetary policy topic (i.e., the second topic) in 2005.
- "Liquidity" first appears in the list in 2007, the year that the financial crisis unfolded.
- "Turmoil", "bubble" and "panic" appear in 2008 and 2009, when the financial crisis was in full swing in the US. What had initially been thought to be a predominantly US problem was turning out to be a global meltdown. During this time the Fed was utilizing all tools at its disposal, including speeches, to try to calm and stabilize the market.

[32] In addition to the Fed, the other Agencies were the Office of the Comptroller of the Currency (OCC), the Federal Deposit Insurance Corporation (FDIC), and the Office of Thrift Supervision (OTS).

- ‘Systemic’ first occurs in 2011 and then again in 2013, in the aftermath of the financial crisis, corresponding to the responses to the Greek crisis and the European sovereign debt crisis (including the closure of Laiki Bank in Cyprus), respectively.
- Despite the stress tests in response to the financial crisis being implemented as early as 2009 (first as the 2009 Supervisory Capital Adequacy Program, or SCAP, and subsequently as the Comprehensive Capital Analysis and Review, or CCAR, that is currently conducted approximately annually), the term “stress_test” does not appear as a top-10 word until 2016.

Interestingly, despite the Fed’s recognition of the importance of communication in effecting policy (see Board of Governors of the Federal Reserve System, 2004), the word “communication” does not appear in the top 10 lists. Finally, notably absent from the word lists as well are the terms “European” and “sovereign”, despite the European sovereign debt crisis that occurred in the aftermath of the 2008-9 financial crisis. Terms related to debt only appear twice, both well prior to the financial crisis (2001 and 2004). It is important to emphasize that this does not mean the Federal Reserve ignored this crisis as no doubt numerous conversations occurred among central bank representatives around the world – only that the topic did not attain prominence in officials’ public speeches.

Figures 15 and 16 show the document-topic proportion estimates, analogous to Figures 12 and 13.[33] Figure 15 displays the average document-topic proportion across all speeches in a

[33] The colors used for the LDA topic interpretations are the same as the colors used for the topic interpretations in the DTM results: economic and business developments is blue, monetary policy is yellow, and risks, regulation and supervision in the banking system is green.

given year. The pattern of each stacked bar is used to indicate the topic number, corresponding to the topic numbers of Tables A1-A20. The color of the stacked bars indicates whether the interpretation of a topic was similar to the interpretation of a topic in other years in the sample. Hence, together with the interpretation of the topics in Table 9, Figure 15 reveals which topics garnered the most attention. For example, from 1998 to 2002, in 2008, and 2016 the speeches paid much attention to the (global) economic and business developments of that time: (in non-chronological order) economic growth, rising energy prices, new technologies in finance, community development programs, or the financial crisis. In 1997, 1998, 1999 and later in 2015 many of the speeches discussed risk management, regulation and supervision of banks and financial institutions. Starting in 2003 and up until 2007 the monetary policy topic received a lot of attention, a result of the approach to, implementation of, and cessation of the tightening cycle. In those years monetary policy was often discussed in relation to inflation, rising energy prices, economic expectations, and (in 2007) housing prices. It is interesting to note that after 2004 the document topic proportions of the monetary policy topic were almost always larger than the economic developments topic (with the exception of 2008 and 2016), while the reverse pattern is observed in the years before. Finally, as a robustness check, the normalized average document-topic proportions and the distribution based on maximum document-topic proportion are displayed in Figures OC4 and OC5 in Online Appendix C. These reveal that the results in Figure 15 are robust to speeches that have very low document-topic proportions (i.e., proportions below 0.05), or when considering only the topic with the most emphasis in each of the speeches.

[Insert Figure 15 here]

Finally, Figure 16 displays the percentage of speeches by year where just one of the three topics was covered extensively. The figure shows that not many speeches were dominated by a single topic. In the years where this was the case, the main topic of the speech was in most cases either monetary policy and inflation (1997-1999, 2000, 2002, 2007, 2012, and 2014) or risks, regulations and supervision of the banking system (2001, 2003, 2013 and 2016). In 2003 and 2004, and in 2012 and 2014 the dominant topic of several speeches was internal control and corporate governance, and financial stability, respectively. The year with the highest percentage of speeches that were dominated by a single topic was 2012; Table 9 and Appendix Table A16 show that those speeches covered financial stability, capital regulations, risks in the financial system, and banking supervision.

[Insert Figure 16 here]

6. Conclusions

Both the DTM model and LDA model results reveal the same three important topics in Federal Reserve speeches: economic and business developments, monetary policy, and risks, regulations, and supervision in the banking system.

The DTM results show a clear separation of the dual mandate of the Federal Reserve among the topics. To the extent that the content of the speeches is indicative of the policy mandate, the economic and business developments topic clearly encapsulated the maximum sustainable employment part of the dual mandate, while the price stability mandate is clearly visible in the top words of the topic on monetary policy. In addition, the document topic proportions reveal how there has been a slight change in the balance of emphasis placed on the two aspects of the dual mandate. Before the financial crisis the Fed was focused more on

creating price stability and stimulating productivity growth, while after the crisis the focus centered more on job creation and moderate longer-term interest rate levels.

In the LDA results the dual mandate is also visible in these two topics, although it is not as clearly separated as in the DTM results. Words related to employment occur not only in the top 10 lists of the economic and business developments topic, but occasionally also in the top 10 lists of the monetary policy topic. The results of the LDA models show that over time, when discussing monetary policy and inflation in speeches, the Fed's emphasis shifted from discussing their relationship to economic growth, investments and rising prices, to discussing their relationship to the federal funds rate, interest rates, and unemployment – this shift began in 2010, coinciding with Bernanke's second term as Chair.

Both the DTM and LDA results revealed that the risks, regulations and supervision of the banking system topic was very important in every year of the sample, and that the language related to this topic (i.e., as defined in the yearly top 10 lists that define the topic) was quite distinct from the language of either of the two mandates in the dual mandate. To the extent that the language of a topic is indicative of the drivers of the Fed's policy stance, these results suggest an important role of the risks, regulations, and supervision of the banking system topic in shaping Fed policy throughout the sample we consider. This is surprising in light of post-crisis assertions that the Fed was not adequately focused on banking supervision prior to the financial crisis (Financial Crisis Inquiry Commission, for example, pages xviii, 54, 94-95).

Yet within this risks, regulations, and supervision of the banking system topic, the results also illustrate the changing nature of the language officials used. Whenever this topic was covered in speeches in the years before the crisis, the focus was on risk management, Basel II, and internal control and compliance -- words that revealed a microprudential approach to

supervision. After the financial crisis, the emphasis shifted towards capital requirements, liquidity, and stress testing -- a change in the language that coincided with a shift from microprudential supervision aimed at monitoring the risk in individual banks, to the more macroprudential focus that has dominated bank supervision post-crisis. Further, the words that define this topic in recent years reveal that the aim of the regulations and supervision in the banking system has been to ensure financial stability. For example, mitigating liquidity risks and preventing another crisis have become important aspects of recent speeches. Similarly, the increase in the probability of words concerning the Fed's asset purchases and the role of the FOMC reveals that more attention is being placed on preventing systemic risks. Taken together, the LDA topic interpretations and the increased importance of financial stability words in the DTM results for the risks, regulations and supervision of the banking system topic highlight the increased emphasis that the Fed has placed on ensuring financial stability, in essence adding this third dimension to their mandate.

This paper has applied two techniques from the field of text analysis to two decades of speeches made by members of the Board of Governors of the Federal Reserve System and has documented a shift in the weights placed on the Fed's dual mandate goals over time. In contrast to previous literature that has sought to address the debate regarding the merits of adopting a third mandate (Peek et al., 2016), our results suggest that financial stability has *de facto* become a third aspect of the Fed's mandate, judging from the amount of rhetoric the Fed has devoted to the topic, consistent with the findings of Oet and Lyytinen (2017).

References

Abe, JoAnn A. (2011), "Changes in Alan Greenspan's Language Use Across the Economic Cycle: A Text Analysis of His Testimonies and Speeches," *Journal of Language and Social Psychology* 30(2), 212-223.

Bailey, A., and C. Schonhardt-Bailey (2008), "Does Deliberation Matter in FOMC Monetary Policymaking? The Volcker Revolution of 1979," *Political Analysis, 16, 404-427.*

Baxter, Thomas C., Jr. (2013), "Financial Stability: The Role of the Federal Reserve System," Remarks at the Future of Banking Regulation and Supervision in the EU Conference, November 15, available at https://www.newyorkfed.org/newsevents/speeches/2013/bax131120, accessed July 20, 2016.

Bernanke, Ben S. (2007), "Federal Reserve Communications," Speech at the Cato Institute 25th Annual Monetary Conference, Washington, D.C., November 14.

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

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

Bligh, Michelle C., and Gregory D. Hess (2007), "The Power of Leading Subtly: Alan Greenspan, Rhetorical Leadership, and Monetary Policy," *The Leadership Quarterly* 18, 87-104.

Blinder, A. S., M. Ehrmann, M. Fratzscher, J. D. Haan, and D.-J. Jansen (2008), "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence," *Journal of Economic Literature*, 46(4), 910–45.

Board of Governors of the Federal Reserve System (2004). *Transcript from the Meeting of the Federal Open Market Committee on March 16, 2004*. Available at www.federalreserve.gov/monetarypolicy/files/FOMC20040316meeting.pdf

Board of Governors of the Federal Reserve System website (source for all speech data), http://www.federalreserve.gov/newsevents/speech/2016speech.htm.

Brainard, Lael (2014), "The Federal Reserve's Financial Stability Agenda," speech given at The Brookings Institution, December 3, available at http://www.federalreserve.gov/newsevents/speech/brainard20141203a.htm, accessed July 20, 2016.

Bullard, James (2010), "The Fed's Dual Mandate: Lessons of the 1970s," Federal Reserve Bank of St. Louis Annual Report, available at https://www.stlouisfed.org/annual-report/2010, accessed August 2, 2016.

Dudley, William C. (2016), "The U.S. Economic Outlook and Implications for Monetary Policy," Remarks at the Economic Leadership Forum, Somerset, New Jersey, January 15, available at https://www.newyorkfed.org/newsevents/speeches/2016/dud160115, accessed August 2, 2016.

Erosheva, E., S. Fienberg, and C. Joutard (2007), "Describing disability through individual-level mixture models for multivariate binary data," *Annals of Applied Statistics*, 1(2):502-537.

Federal Open Market Committee (2014). *Transcript from the Meeting of the Federal Open Market Committee*. March 16, May 4, and June 29-30, available respectively at https://www.federalreserve.gov/monetarypolicy/files/FOMC20040316meeting.pdf; https://www.federalreserve.gov/monetarypolicy/files/FOMC20040504meeting.pdf, and https://www.federalreserve.gov/monetarypolicy/files/FOMC20040630meeting.pdf

Fei-Fei, Li, and Pietro Perona (2005), "A Bayesian Hierarchical Model for Learning Natural Scene Categories," Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2: 524-531.

Financial Crisis Inquiry Commission (2011). *Financial Crisis Inquiry Report.* Washington, DC: US Government Printing Office. Available at https://www.govinfo.gov/content/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf.

Fligstein, N., J.S., Brundage, and M. Schultz (2014), "Why the Federal Reserve Failed to See the Financial Crisis of 2008: The Role of "Macroeconomics" as a Sensemaking and Cultural Frame," Mimeograph, University of California Berkeley.

Grimmer, J. (2011), "An Introduction to Bayesian Inference via Variational Approximations," *Political Analysis* 19(1), 32-47.

Haltom, Renee, and Jeffrey M. Lacker (2013), "Should the Fed Have a Financial Stability Mandate: Lessons from the Fed's First 100 Years," Federal Reserve Bank of Richmond Annual Report, available at https://www.richmondfed.org/~/media/richmondfedorg/publications/research/annual_report/2013/pdf/article.pdf, accessed July 20, 2016

Hansen, Stephen, and Michael McMahon (2016), "Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication," *Journal of International Economics*, 99(1): S114-S133.

Hansen, Stephen, Michael McMahon, and Andrea Prat (2014), "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach," CEP Discussion Papers dp1276, Centre for Economic Performance, LSE.

Jiang S, Qian X, Shen J, Fu Y, Mei T. (2015), "Author topic model-based collaborative filtering for personalized POI recommendations," IEEE Trans Multimedia,17(6):907–918.

Kahn, George A. and Lisa Taylor (2014), "Evolving Market Perceptions of Federal Reserve Policy Objectives," Federal Reserve Bank of Kansas City *Economic Review* (first quarter).

Kohn, Donald L. and Brian P. Sack (2004), "Central Bank Talk: Does It Matter and Why?", in Bank of Canada (ed.), *Macroeconomics, Monetary Policy, and Financial Stability*, pp. 175-206, Ottawa: Bank of Canada.

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

Lucca, David O. and Francesco Trebbi (2009), "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements," NBER working paper 15367.

Mester, Loretta J. (2016), "Monetary Policy and Financial Stability in the U.S.," speech delivered at the Sydney Banking and Financial Stability Conference, University of Sydney, July 12, available at https://www.clevelandfed.org/newsroom-and-events/speeches/sp-20160712-monetary-policy-and-financial-stability-in-the-us.aspx, accessed July 15, 2016.

Mandt, Stephan, and Blei, David M. (2014) "Smoothed Gradients for Stochastic Variational Inference." Neural Information Processing Systems 2014: 2438-2446.

Michel, Norbert (2014), "Dodd-Frank's Expansion of Fed Power: A Historical Perspective," *Cato Journal* 34(3), 557-567.

Mikolov, T., I. Sutskever, K. Chen, G. Corrado, and J. Dean (2013), 'Distributed Representations of Words and Phrases and their Compositionality,' Proceedings of the Neural and Information Processing Systems Conference (NIPS) 2013.

Mimno, D., H.M. Wallach, H.M., E. Talley, M. Leenders, and A. McCallum (2011), 'Optimizing Semantic Coherence in Topic Models', Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, 262-272.

Oet, M.V., and Lyytinen, K. (2017) Does Financial Stability Matter to the Fed in Setting US Monetary Policy? *Review of Finance* 21(1): 389-432.

Peek, Joe, Eric S. Rosengren, and Geoffrey M.B. Tootell (2016), "Does Fed Policy Reveal a Ternary Mandate?" *Working Papers 16-11, Federal Reserve Bank of Boston*.

Pianalto, Sandra (2011), "Maximum Employment: What We Know (and Don't Know) about the Labor Market," December 31, available at https://www.clevelandfed.org/en/newsroom-and-events/publications/annual-reports/ar-2011-maximum-employment/ar-201102-maximum-employment.aspx, accessed August 2, 2016.

Ramage, Daniel, Susan Dumais, and Dan Liebling (2010), 'Characterizing microblogs with topic models," Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 130-137.

Resche, Catherine (2004), "Investigating 'Greenspanese': From Hedging to 'Fuzzy Transparency'," *Discourse & Society* 15(6), 723-744.

Resche, Catherine (2015), "Hedging in the Discourse of Central Banks," *Studies in Communication Sciences* 15, 83-92.

Schonhardt-Bailey, C. (2013). *Deliberating Monetary Policy*. Cambridge, MA: MIT Press.

Srivastava, Ahok, N. and Mehran Sahami (2009). *Text Mining: Classification, Clustering, and Applications.* CRC Press.

Steelman, Aaron (2011), "The Federal Reserve's 'Dual Mandate': The Evolution of an Idea," Federal Reserve Bank of Richmond Economic Brief #EB11-12, available at https://www.richmondfed.org/~/media/richmondfedorg/publications/research/economic_brief/2011/pdf/eb_11-12.pdf, accessed July 20, 2016.

Steyvers, M., and Griffiths, T. (2006). Probabilistic Topic Models. In: *Latent Semantic Analysis: A Road to Meaning*. T. Landauer, D. McNamara, S. Dennis, and W. Kintsch, eds. Laurence Erlbaum, 2006.

Williams, John C. (2012), "The Federal Reserve's Mandate and Best Practice Monetary Policy," Presentation to the Marian Miner Cook Athenaeum, Claremont McKenna College, February 13, available at http://www.frbsf.org/our-district/press/presidents-speeches/williams-speeches/2012/february/williams-federal-reserve-mandate-best-practice-monetary-policy/, accessed July 20, 2016.

FIGURES

Figure 1 – The number of speeches per year.

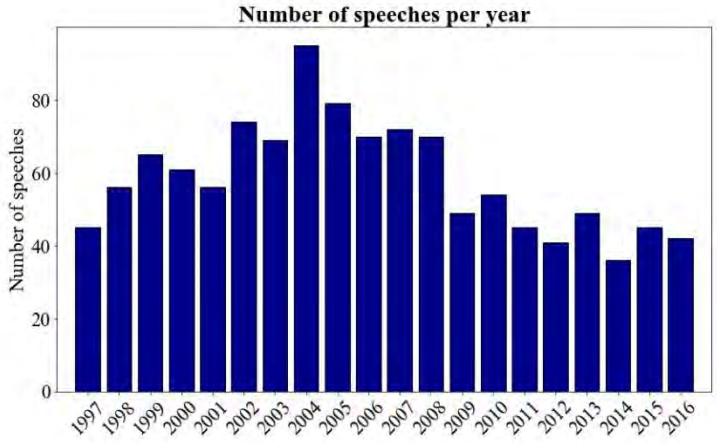


Figure 2 – Number of types and tokens per speech over time. The line represents the least squares polynomial fit.

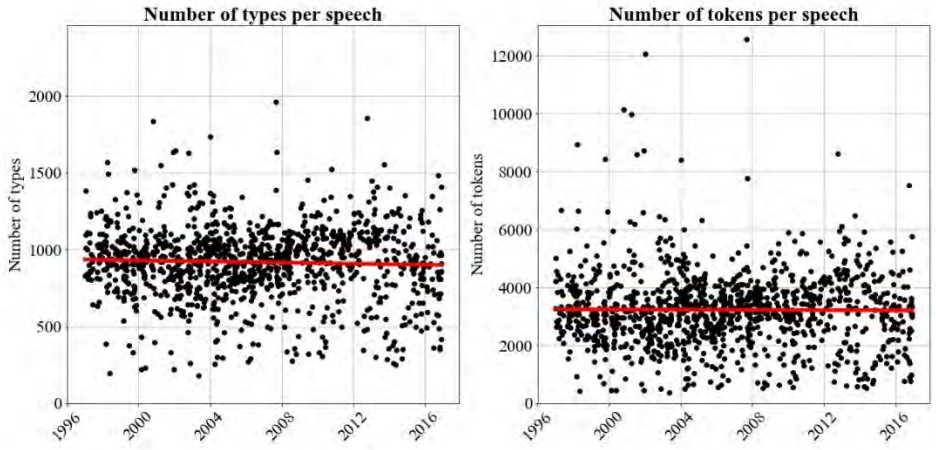


Figure 3 – Number of types and tokens per speech. The line represents the least squares polynomial fit.

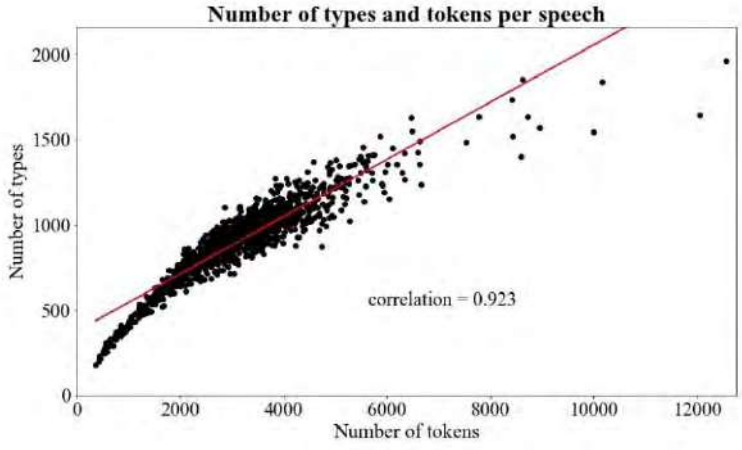


Figure 4 – Mandate Dictionary results, expressed as the average percentage of total number of tokens per speech per year

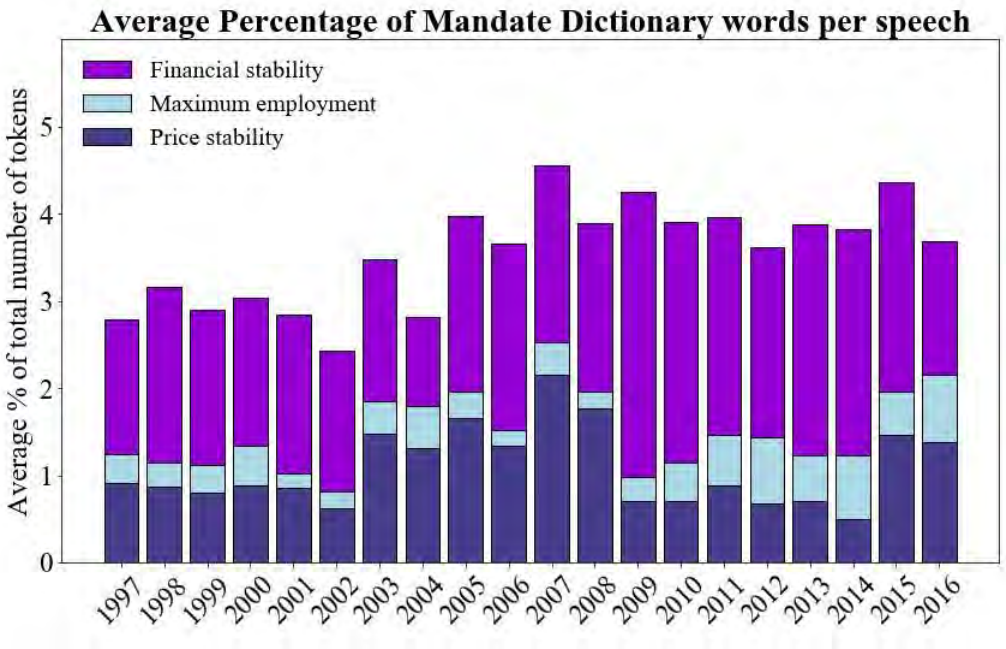


Figure 5 – Mandate Dictionary results with Peek et al. (2016) financial stability words, expressed as the average percentage of total number of tokens per speech per year

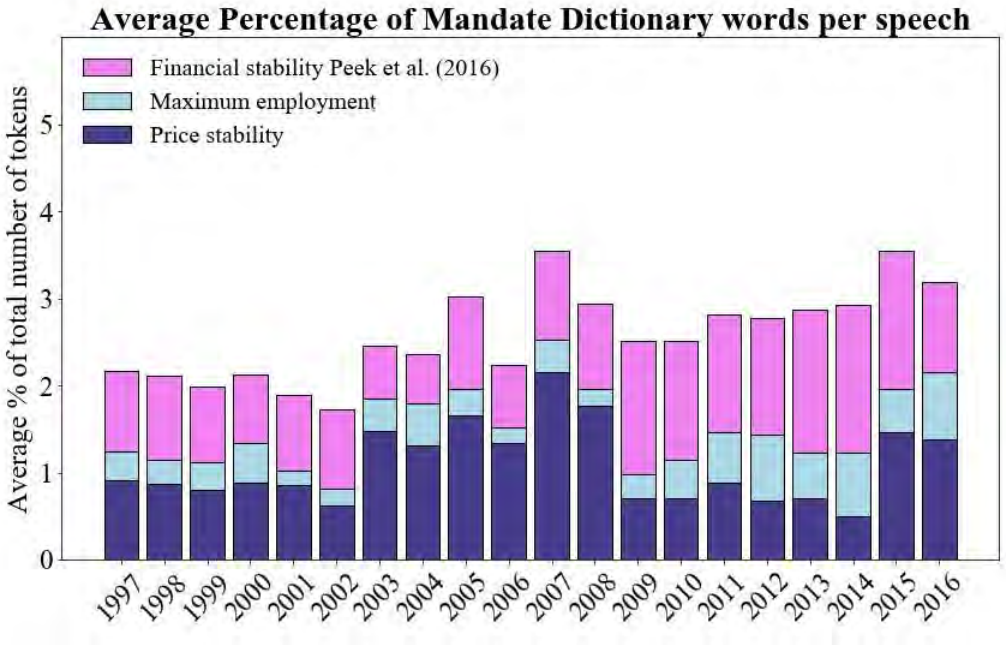


Figure 6 – Graphical model representation of LDA, as first presented in Blei et al., 2003. The notation of the original graphical model is slightly adapted to make it consistent with the notation employed in this paper.

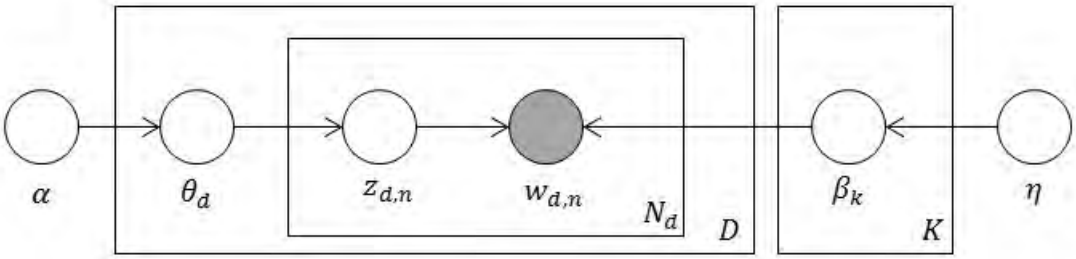


Figure 7 – Graphical model representation of DTM (Blei et al., 2006). The notation of the original graphical model is slightly adapted to make it consistent with the notation employed in this paper. Each column represents a time slice $\boldsymbol{t}, \boldsymbol{t} = \mathbf{1}$ ....**T**

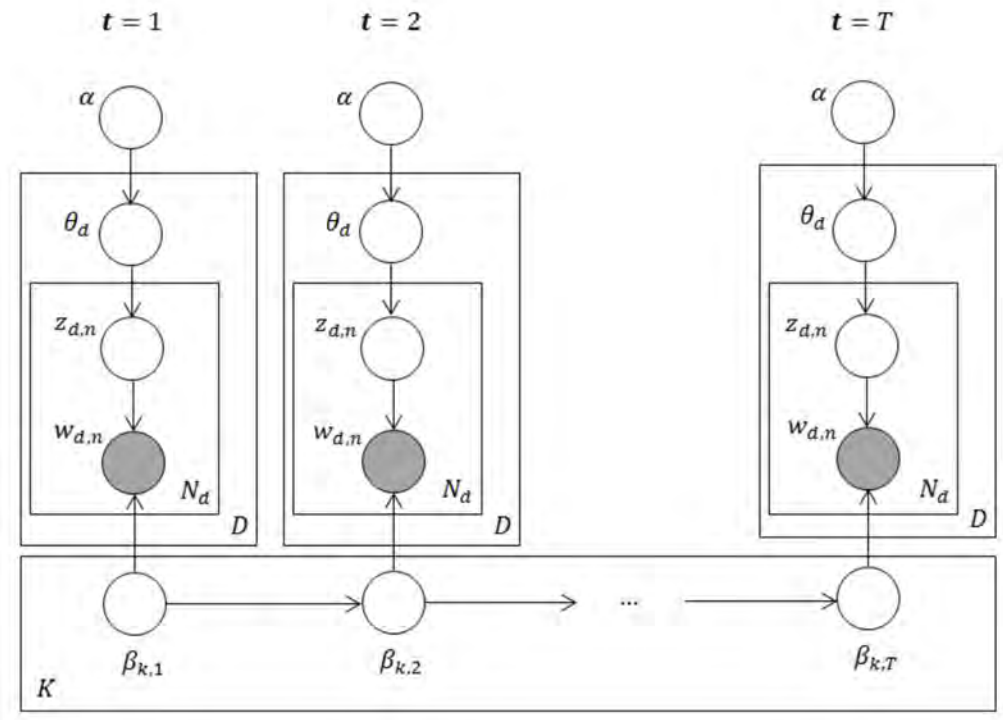


Figure 8 – Evolution of the term scores of the words that occurred at least once in the top 5 of topic 1 from 1997 to 2016.

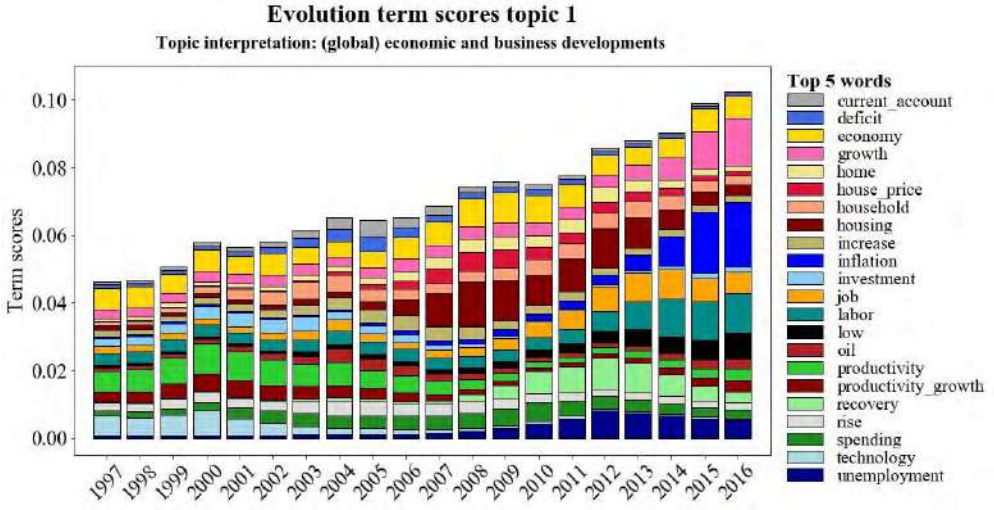


Figure 9 – Evolution of the term scores of the words that occurred at least once in the top 5 of topic 2 from 1997 to 2016.

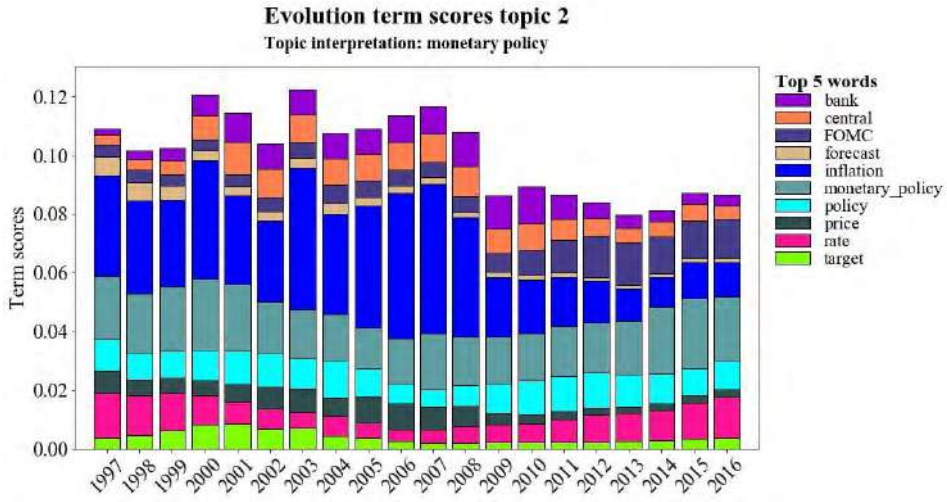


Figure 10 – Evolution of the term scores of the words that occurred at least once in the top 5 of topic 3 from 1997 to 2016.

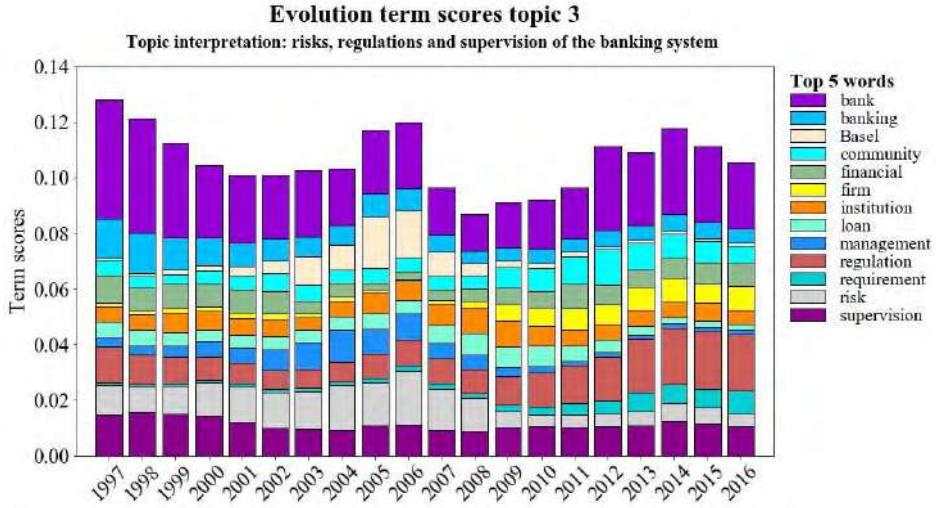


Figure 11 – The word clouds of the top 100 words of the three topics before, during, and after the crisis, as estimated by the DTM. The size of each word is indicative of the probability and the color shows the topic in which it obtained the highest probability. The left column contains word clouds based on the word probabilities and the right column contains the associated word clouds based on the termscores.

Figure 12 –The average document-topic proportions across all speeches in a given year, from 1997 to 2016, as estimated by the DTM.

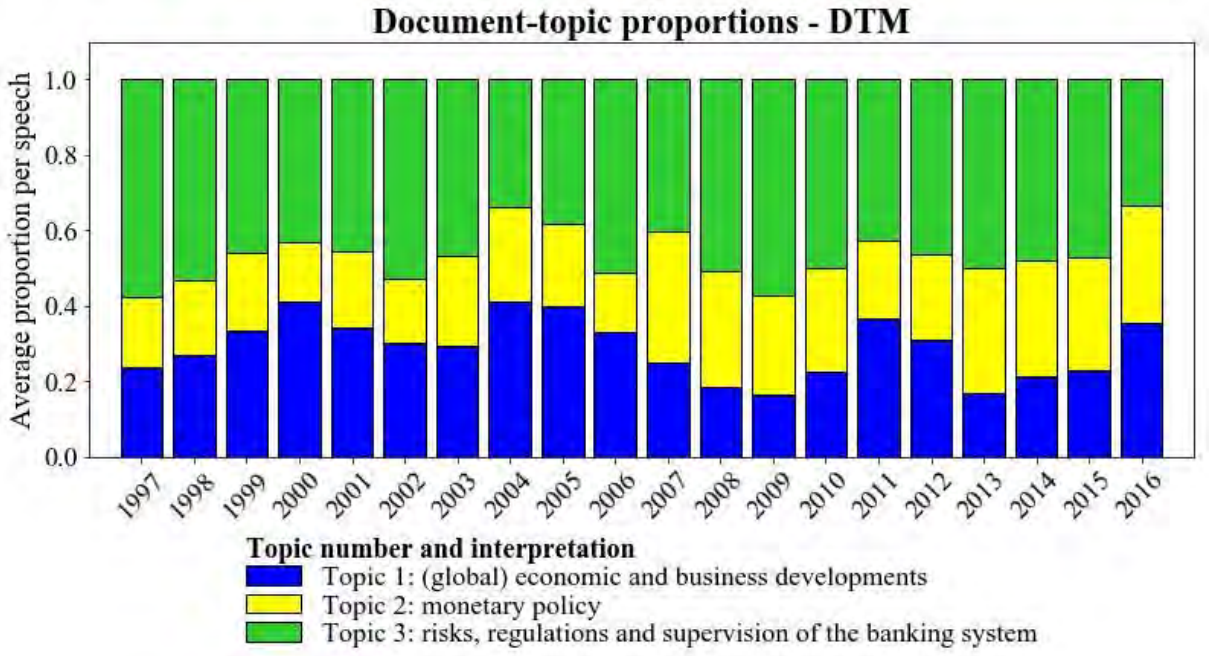


Figure 13 –The percentage of speeches dominated by a single topic (as measured by a document-topic proportion above 0.95), by year, as estimated by the DTM.

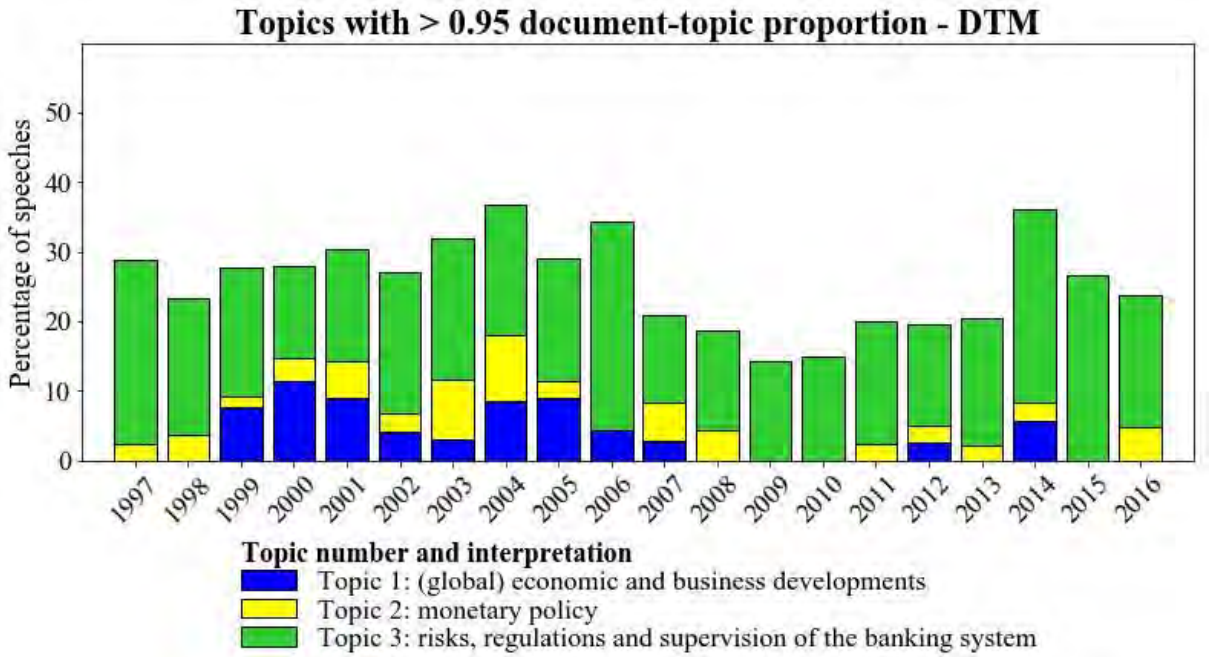


Figure 14 – The average percentage of times the word ‘inflation’ preceded the word ‘expectation’, calculated per year from 1997 to 2016. For reference, the annual inflation rate (as measured by the year-on-year change in the Consumer Price Index, blue line) and the federal funds rate (orange line) are also shown.

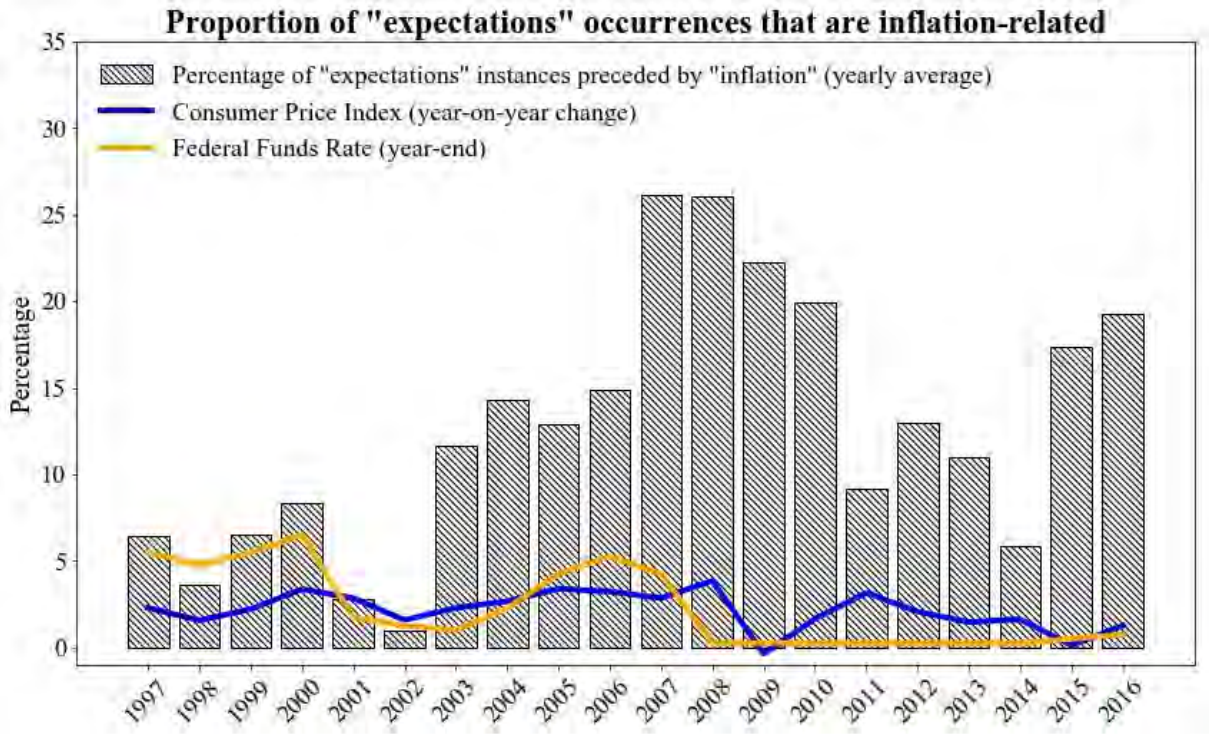


Figure 15 – The average document-topic proportions across all speeches in a given year, from 1997 to 2016, as estimated by the LDA model.

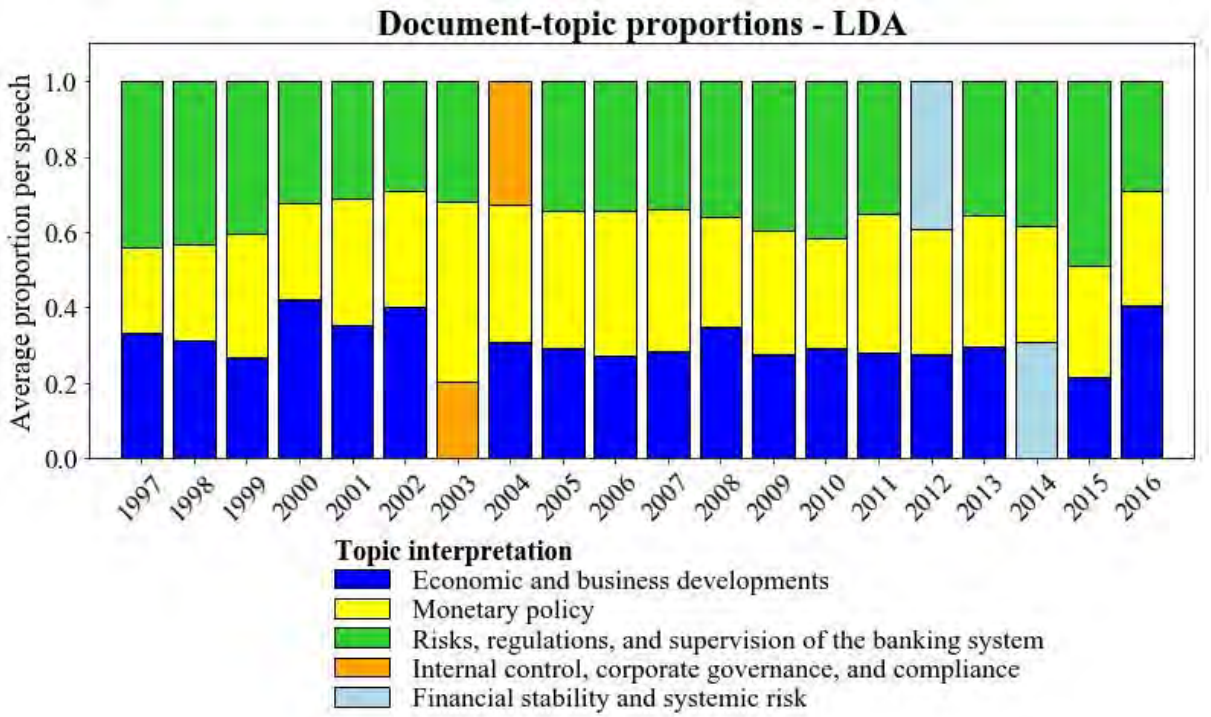
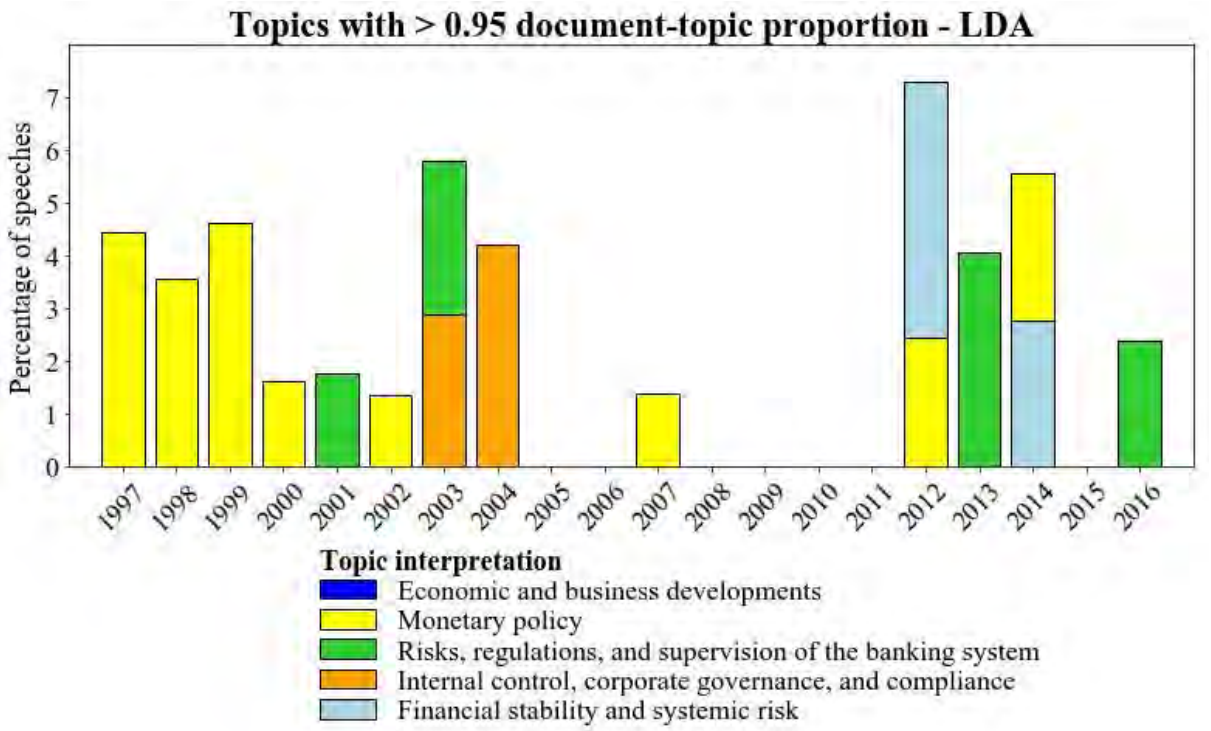


Figure 16 – The percentage of speeches dominated by a single topic (as measured by document-topic proportion above 0.95), by year, as estimated by the LDA model.



TABLES

Table 1 –Federal Reserve Governors from January 1997 (start of sample) to December 2016 (end of sample), with the corresponding number of speeches, tokens and types.[34]

Name	Position	Dates of Term	# of speeches in sample	# of tokens	# of types	Average TTR[35]
Alan Greenspan	Chair	Start of sample – January 31, 2006	168	442,773	11,824	0.35
Ben Bernanke	Governor	August 5, 2002 – January 31, 2006	40	173,101	8,170	0.27
	Chair	February 1, 2006 – January 31, 2014	148	455,058	11,141	0.32
Janet Yellen*	Vice-Chair	October 4, 2010 – February 2, 2014	22	84,101	5,386	0.28
	Chair	February 3, 2014 – end of sample	23	68,435	4,436	0.30
Alice Rivlin	Vice-Chair	Start of sample – July 16, 1999	15	40,340	4,886	0.33
Roger Ferguson	Governor	November 5, 1997 – October 4, 1999	27	78,790	5,522	0.31
	Vice-Chair	October 5, 1999 – April 28, 2006	91	293,033	9,426	0.30
Donald Kohn	Governor	August 5, 2002 – June 22, 2006	29	94,758	5,855	0.29
	Vice-Chair	June 23, 2006 – September 1, 2010	45	136,267	6,466	0.31
Stanley Fischer*	Vice-Chair	June 16, 2014 – end of sample	35	94,811	5,734	0.31
Edward Kelley	Governor	Start of sample – December 31, 2001	8	21,241	3,128	0.33
Susan Phillips	Governor	Start of sample – June 30, 1998	10	29,406	3,326	0.30
Laurence Meyer	Governor	Start of sample – January 31, 2002	71	329,739	9,229	0.25
Edward Gramlich	Governor	November 5, 1997 – August 31, 2005	57	161,014	7,946	0.31
Mark Olson	Governor	December 7, 2001 – June 30, 2006	44	112,603	6,547	0.32
Susan Bies	Governor	December 7, 2001 – March 30, 2007	77	263,171	7,077	0.28
Kevin Warsh	Governor	February 24, 2006 – April 2, 2011	18	49,845	5,263	0.35
Randall Kroszner	Governor	March 1, 2006 – January 21, 2009	43	139,651	6,536	0.28
Frederic Mishkin	Governor	September 5, 2006 – August 31, 2008	28	114,442	6,187	0.26
Elizabeth Duke	Governor	August 5, 2008 – August 31, 2013	42	149,313	7,046	0.29
Daniel Tarullo*	Governor	January 28, 2009 – end of sample	51	215,124	7,663	0.28
Sarah Raskin	Governor	October 4, 2010 – March 13, 2014	17	60,495	5,355	0.30
Jerome Powell*	Governor	May 25, 2012 – end of sample	32	79,779	5,417	0.32
Jeremy Stein	Governor	May 30, 2012 – May 28, 2014	15	50,345	4,130	0.29
Lael Brainard*	Governor	June 16, 2014 – end of sample	17	50,855	4,401	0.29

[34] The full dates that each member served are available in https://www.federalreserve.gov/aboutthefed/bios/board/boardmembership.htm. In this table we only focus on the dates that each served as they relate to our sample. Specifically, the positions of people that did not give any speeches during the time period of our sample are not included in the list (i.e., the list does not include Lawrence Lindsey and the governorships of Janet Yellen and Stanley Fischer). Members that were still in office as of the end of the data sample are denoted with an asterisk (*).

[35] For each speech the TTR was calculated, and the average TTR was calculated per person over all the speeches they gave during their respective date of term as Governor, Vice-Chair or Chair.

Table 2 – The twelve speeches that rank in the Top 10 either by number of tokens or number of types, listed according to the number of tokens. For each speech its ranking by token (column 6) and type (column 7), as well as the corresponding number of tokens, types, and type-to-token ratio (TTR) is provided.

Governor	Date	# of tokens	# of types	TTR	Ranking in # of tokens	Ranking in # of types
Frederic Mishkin	2007-09-01	12,557	1,960	0.16	1	1
Laurence Meyer	2002-01-16	12,053	1,642	0.14	2	5
Laurence Meyer	2000-10-24	10,154	1,836	0.18	3	3
Laurence Meyer	2001-03-28	9,988	1,547	0.15	4	11
Laurence Meyer	1998-04-02	8,953	1,568	0.18	5	9
Laurence Meyer	2001-12-05	8,727	1,634	0.19	6	7
Daniel Tarullo	2012-10-10	8,625	1,855	0.22	7	2
Laurence Meyer	2001-07-17	8,590	1,401	0.16	8	24
Laurence Meyer	1999-10-12	8,434	1,520	0.18	9	13
Roger Ferguson	2004-01-04	8,404	1,733	0.21	10	4
Frederic Mishkin	2007-09-21	7,779	1,635	0.21	11	6
Ben Bernanke	2002-10-15	6,463	1,628	0.25	18	8
Daniel Tarullo	2013-09-20	6,479	1,552	0.24	17	10

Table 3 – Top 10 list of the speeches according to the type-to-token ratio (TTR). For each speech the corresponding number of tokens, types, and type-to-token ratio (TTR) is given.

Governor	Date	# of tokens	# of types	TTR	Ranking in top 10 TTR
Ben Bernanke	2010-04-13	562	308	0.55	1
Alan Greenspan	2000-05-18	443	232	0.52	2
Kevin Warsh	2008-07-28	534	278	0.52	3
Alan Greenspan	2002-11-18	505	259	0.51	4
Ben Bernanke	2008-10-14	633	322	0.51	5
Ben Bernanke	2004-04-01	501	254	0.51	6
Alan Greenspan	2001-10-23	659	332	0.50	7
Ben Bernanke	2013-11-14	576	289	0.50	8
Ben Bernanke	2009-06-26	642	322	0.50	9
Alan Greenspan	2003-05-13	361	180	0.50	10

Table 4 – Categories of the Mandate Dictionary and the words (not case-sensitive) associated with those three categories. As a robustness check, two different sets of words are used for the financial stability category, one is self-constructed (as the other categories are) and the other is based on the words from the Peek et al. (2016) paper.[36]

Mandate Dictionary Category	Words
price stability	inflation, inflationary, deflation, deflationary, price, price stability, cpi, pce, price index
maximum employment	employment, unemployment, job, nairu, displaced worker, payroll, layoff, hire, continuing claim, establishment level
financial stability	basel, supervisor, supervision, capital, credit, credit crisis, credit crunch, systemic, systemic risk, panic, crisis, financial crisis, global financial crisis, bubble, burst, bust, collapse, crash, financial stability, financial instability
financial stability Peek et al. (2016)	anxiety, asset price, bubble, burst, bust, cds, collapse, crash, credit constrain, crisis, crunch, equity price, equity, equity value, financial stability, froth, house price, housing price, illiquidity, instability, irrational exuberance, ldc, lending standard, liquidity issue, liquidity problem, loan officer, market correction, market distress, panic, price to earnings, regulation, stock market, stock price, supervision, volatility

[36] As the search is conducted over the texts after lemmatization and bigram detection, the dictionary words are adjusted accordingly. For example, there is no separate search for 'crisis' and 'crises', but just for 'crisis'). Moreover, instead of manually checking beforehand whether the collocations in the dictionary are connected via an underscore in the texts or not, the search is conducted over both forms. For example, both 'price index' and 'price_index' are counted, and because one of the two occurs zero times, the total count per dictionary category is the same as it would be if we would have manually checked beforehand for each collocation. Instead of listing both forms of the phrase, Table 4 just lists the phrases with white space in between the words. Finally, as the modified texts do not contain stopwords and only contain bigrams and not trigrams, the term 'price to earnings' would not have been included. In order to include the counts of that phrase for comparison with the Peek et al. (2016) results, a copy of the set of speeches has been created in which the phrase 'price to earnings' was turned into the trigram 'price_to_earnings' after lemmatization and before continuing with stopwords removal. This set of texts was only used to conduct the search using the dictionary with the Peek et al. (2016) financial stability category, in order to match their list as closely as possible.

Table 5 – Terms and their respective meaning employed in the LDA and DTM models. For the DTM model, a subscript t is added to indicate that the parameter concerns a time slice (is time-varying) as opposed to being constant over the entire corpus

Term	Meaning
D	The number of documents in the corpus, indexed by $d = 1, \dots, D$.
V	The number of words in the vocabulary, indexed by $v = 1, \dots, V$.
K	The number of topics, each indexed by $k = 1, \dots, K$.
N_d	The number of words in the dth document, indexed by $n = 1, \dots, N_d$.
$w_{d,n}$	The nth word in the dth document. It is a V-dimensional unit vector with components $w_{d,n}^v, v = 1, \dots, V$.
$\boldsymbol{w}_d$	The dth document, containing N_d words: $\boldsymbol{w}_d = (w_{d,1}, w_{d,2}, \dots., w_{d,N_d})$.
$\boldsymbol{w}_{1:D}$	The corpus of D documents: $\boldsymbol{w}_{1:D} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_D\}$.
β_k	The kth topic (i.e., the kth topic-word distribution). It is a V-dimensional vector of probabilities.
$\beta_{k,v}$	The probability that the vth word in the vocabulary occurs in topic β_k's distribution (taken over the number of words in the vocabulary).
β	The $K \times V$ matrix of all the topics of the corpus, where each row corresponds to a topic $\beta_k, k = 1, \dots, K$.
θ_d	The document-topic distribution of the dth document ($\boldsymbol{w}_d$). It is a K-dimensional vector of probabilities.
$\theta_{d,k}$	The probability that the kth topic occurs in the document-topic distribution (θ_d) of the dth document ($\boldsymbol{w}_d$).
θ	The $D \times K$ matrix of all the document-topic distributions of the corpus, where each row corresponds to a document-topic distribution $\theta_d, d = 1, \dots, D$.
η_v	The prior weight of the vth word of the vocabulary in topic β_k.
η	The V-dimensional vector of prior weights η_v, $v = 1, \dots, V$ for all the V words in the vocabulary -- also referred to as the *hyperparameter of the topic-word distribution*.
α_k	The prior weight of the kth topic (β_k) in document $\boldsymbol{w}_d$.
α	The K-dimensional vector of prior weights α_k, $k = 1, \dots, K$ for all the K topics -- also referred to as the *hyperparameter of the document-topic distribution.*
$z_{d,n}$	The topic that is assigned to the nth word in the dth document ($w_{d,n}$). It is a K-dimensional unit vector with components $z_{d,n}^k$, $k = 1, \dots, K$.
$\boldsymbol{z}_d$	The topic assignments of the N_d number of words in each document $\boldsymbol{w}_d$: $\boldsymbol{z}_d = (z_{d,1}, z_{d,2}, \dots., z_{d,N_d})$.
$\boldsymbol{z}_{1:D}$	The collection of topic assignments of the corpus: $\boldsymbol{z}_{1:D} = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_D\}$.
$\beta_{z_{d,n}}$	The topic-word distribution corresponding to the value of the topic-assignment $z_{d,n}$.

Table 6a –Top 10 words of topic 1 per year (by topic-word probability), estimated using the DTM

Topic 1 – Word probabilities top 10									
Words_1997	**Words_1998**	**Words_1999**	**Words_2000**	**Words_2001**	**Words_2002**	**Words_2003**	**Words_2004**	**Words_2005**	**Words_2006**
economy	economy	economy	economy	economy	economy	economy	economy	economy	economy
increase	increase	growth	increase	investment	investment	increase	increase	price	increase
growth	growth	increase	growth	increase	increase	investment	rate	increase	price
new	new	new	investment	capital	rate	rate	price	rate	rate
price	price	capital	new	growth	growth	growth	rise	investment	growth
technology	investment	investment	capital	rate	capital	price	investment	rise	investment
high	high	technology	technology	price	business	rise	growth	growth	high
investment	capital	rate	rate	new	price	business	household	high	rise
change	technology	high	high	high	high	capital	high	recent	recent
work	rate	price	business	business	rise	household	business	low	low
Words_2007	**Words_2008**	**Words_2009**	**Words_2010**	**Words_2011**	**Words_2012**	**Words_2013**	**Words_2014**	**Words_2015**	**Words_2016**
economy	economy	economy	economy	economy	economy	economy	economy	economy	economy
increase	increase	increase	rate	recovery	recovery	rate	rate	inflation	growth
rate	rate	rate	increase	increase	rate	low	growth	growth	inflation
price	housing	housing	growth	rate	housing	recovery	low	rate	low
growth	price	growth	recovery	growth	low	growth	inflation	low	rate
high	growth	high	housing	housing	increase	increase	recovery	increase	increase
housing	high	price	high	low	growth	job	increase	labor	labor
investment	rise	household	household	high	unemployment	work	labor	decline	decline
rise	recent	rise	low	household	mortgage	housing	job	work	work
recent	investment	low	business	mortgage	household	household	work	price	recent

Table 6b –Top 10 words of topic 1 per year (by term score), estimated using the DTM

Topic 1 – Term scores top 10									
Words_1997	**Words_1998**	**Words_1999**	**Words_2000**	**Words_2001**	**Words_2002**	**Words_2003**	**Words_2004**	**Words_2005**	**Words_2006**
economy	productivity	productivity	productivity	productivity	productivity	productivity	productivity	productivity	economy
productivity	economy	technology	technology	economy	economy	economy	household	current_account	productivity
technology	technology	economy	economy	productivity_growth	productivity_growth	household	rise	economy	housing
labor	labor	productivity_growth	productivity_growth	technology	investment	investment	economy	deficit	increase
productivity_growth	productivity_growth	labor	investment	investment	household	spending	oil	rise	spending
new	new	growth	labor	spending	spending	productivity_growth	labor	increase	rate
growth	growth	rise	growth	household	growth	rise	deficit	spending	growth
product	investment	social_security	business	capital	technology	labor	spending	labor	energy
business	rise	new	rise	growth	business	growth	current_account	household	labor
production	business	business	capital	labor	labor	business	increase	rate	rise
Words_2007	**Words_2008**	**Words_2009**	**Words_2010**	**Words_2011**	**Words_2012**	**Words_2013**	**Words_2014**	**Words_2015**	**Words_2016**
housing	housing	housing	housing	housing	housing	housing	labor	inflation	inflation
economy	economy	economy	economy	recovery	recovery	recovery	inflation	labor	growth
house_price	house_price	house_price	recovery	economy	unemployment	job	job	growth	labor
increase	spending	spending	spending	unemployment	job	labor	growth	economy	low
spending	home	home	job	job	economy	unemployment	unemployment	job	economy
productivity	growth	recovery	household	spending	labor	income	recovery	low	job
rate	increase	growth	unemployment	household	mortgage	economy	housing	rate	rate
growth	rise	household	fiscal	labor	household	low	economy	unemployment	unemployment
labor	labor	job	growth	income	home	household	low	recovery	decline
rise	rate	labor	house_price	fiscal	income	inflation	income	recession	labor_force

Table 7a – Top 10 words of topic 2 per year (by topic-word probability), estimated using the DTM

Topic 2 – Word probabilities top 10									
Words_1997	**Words_1998**	**Words_1999**	**Words_2000**	**Words_2001**	**Words_2002**	**Words_2003**	**Words_2004**	**Words_2005**	**Words_2006**
inflation	inflation	rate	inflation	inflation	inflation	inflation	inflation	inflation	inflation
rate	rate	inflation	rate	policy	policy	policy	policy	price	price
policy	economy	economy	policy	rate	rate	price	economy	economy	economy
economy	policy	policy	monetary_policy	monetary_policy	price	economy	rate	policy	policy
price	monetary_policy	monetary_policy	economy	bank	economy	rate	price	rate	financial
monetary_policy	price	price	price	economy	bank	bank	bank	bank	bank
growth	growth	growth	bank	price	monetary_policy	monetary_policy	monetary_policy	central	rate
interest_rate	interest_rate	bank	central	central	central	central	central	monetary_policy	monetary_policy
change	country	country	growth	target	interest_rate	target	financial	financial	central
exchange	change	interest_rate	target	growth	target	interest_rate	expectation	expectation	Federal_Reserve
Words_2007	**Words_2008**	**Words_2009**	**Words_2010**	**Words_2011**	**Words_2012**	**Words_2013**	**Words_2014**	**Words_2015**	**Words_2016**
inflation	inflation	economy	economy	policy	policy	policy	policy	rate	rate
economy	economy	financial	policy	economy	economy	economy	economy	economy	economy
price	financial	policy	financial	financial	rate	rate	rate	policy	policy
financial	price	inflation	bank	rate	financial	financial	monetary_policy	monetary_policy	monetary_policy
monetary_policy	bank	bank	inflation	inflation	monetary_policy	monetary_policy	financial	financial	financial
bank	policy	rate	rate	monetary_policy	inflation	Federal_Reserve	inflation	inflation	inflation
policy	rate	Federal_Reserve	Federal_Reserve	Federal_Reserve	Federal_Reserve	inflation	interest_rate	interest_rate	interest_rate
rate	monetary_policy	monetary_policy	monetary_policy	bank	bank	FOMC	fund	fund	fund
central	central	price	central	central	FOMC	asset	Federal_Reserve	FOMC	price
Federal_Reserve	Federal_Reserve	central	price	price	central	committee	FOMC	price	FOMC

Table 7b – Top 10 words of topic 2 per year (by term score), estimated using the DTM

Topic 2 – Term scores top 10									
Words_1997	**Words_1998**	**Words_1999**	**Words_2000**	**Words_2001**	**Words_2002**	**Words_2003**	**Words_2004**	**Words_2005**	**Words_2006**
inflation	inflation	inflation	inflation	inflation	inflation	inflation	inflation	inflation	inflation
monetary_policy	monetary_policy	monetary_policy	monetary_policy	monetary_policy	monetary_policy	monetary_policy	monetary_policy	monetary_policy	monetary_policy
rate	rate	rate	policy	policy	policy	policy	policy	policy	central
policy	policy	policy	rate	central	central	central	central	price	price
price	forecast	target	target	bank	bank	bank	bank	central	bank
forecast	price	price	central	target	price	price	rate	bank	policy
interest_rate	target	central	bank	rate	rate	target	FOMC	FOMC	FOMC
exchange	interest_rate	forecast	price	price	target	rate	price	rate	expectation
growth	FOMC	bank	monetary	money	monetary	FOMC	monetary	economy	economy
monetary	growth	monetary	FOMC	objective	FOMC	monetary	economy	monetary	stability
Words_2007	**Words_2008**	**Words_2009**	**Words_2010**	**Words_2011**	**Words_2012**	**Words_2013**	**Words_2014**	**Words_2015**	**Words_2016**
inflation	inflation	inflation	inflation	monetary_policy	monetary_policy	monetary_policy	monetary_policy	monetary_policy	monetary_policy
monetary_policy	monetary_policy	monetary_policy	monetary_policy	inflation	inflation	FOMC	FOMC	FOMC	rate
central	bank	bank	bank	policy	FOMC	inflation	policy	rate	FOMC
bank	central	policy	policy	FOMC	policy	policy	rate	inflation	inflation
price	price	central	central	bank	rate	rate	inflation	policy	policy
policy	policy	FOMC	FOMC	rate	central	committee	committee	interest_rate	interest_rate
FOMC	rate	rate	rate	central	committee	central	central	central	committee
economy	economy	economy	economy	economy	bank	bank	interest_rate	Fed	economy
expectation	FOMC	Federal_Reserve	Federal_Reserve	committee	economy	interest_rate	Fed	committee	Fed
rate	expectation	price	interest_rate	Federal_Reserve	interest_rate	economy	economy	economy	central

Table 8a – Top 10 words of topic 3 per year (by topic-word probability), estimated using the DTM

Topic 3 – Word probabilities top 10									
Words_1997	**Words_1998**	**Words_1999**	**Words_2000**	**Words_2001**	**Words_2002**	**Words_2003**	**Words_2004**	**Words_2005**	**Words_2006**
bank	bank	bank	bank	bank	bank	bank	risk	risk	risk
financial	financial	financial	risk	risk	risk	risk	bank	bank	bank
risk	risk	risk	financial	financial	financial	financial	financial	financial	financial
banking	banking	banking	system	system	management	management	management	management	management
regulation	system	system	banking	capital	credit	credit	credit	institution	capital
system	regulation	institution	institution	institution	institution	capital	institution	Basel	institution
supervision	supervision	supervision	supervision	banking	banking	banking	banking	capital	regulation
institution	capital	regulation	capital	supervision	system	regulation	regulation	credit	Basel
loan	institution	capital	regulation	credit	regulation	institution	new	regulation	banking
Federal_Reserve	loan	credit	management	management	supervision	supervision	loan	banking	credit
Words_2007	**Words_2008**	**Words_2009**	**Words_2010**	**Words_2011**	**Words_2012**	**Words_2013**	**Words_2014**	**Words_2015**	**Words_2016**
risk	risk	financial	financial	financial	bank	bank	bank	bank	financial
bank	financial	bank	bank	bank	financial	financial	financial	financial	bank
financial	bank	risk	risk	firm	regulation	regulation	regulation	regulation	regulation
credit	credit	credit	firm	risk	risk	risk	risk	risk	firm
institution	institution	institution	regulation	regulation	firm	firm	firm	firm	capital
regulation	mortgage	firm	credit	capital	community	capital	capital	capital	risk
loan	loan	regulation	system	system	capital	community	system	system	system
lender	lender	loan	institution	community	system	institution	large	institution	liquidity
mortgage	regulation	lender	capital	institution	institution	system	supervision	large	large
capital	capital	capital	loan	lender	lender	large	community	liquidity	institution

Table 8b – Top 10 words of topic 3 per year (by term score), estimated using the DTM

Topic 3 – Term scores top 10									
Words_1997	**Words_1998**	**Words_1999**	**Words_2000**	**Words_2001**	**Words_2002**	**Words_2003**	**Words_2004**	**Words_2005**	**Words_2006**
bank	bank	bank	bank	bank	bank	bank	bank	bank	bank
supervision	supervision	supervision	supervision	risk	risk	risk	risk	Basel	risk
banking	banking	banking	risk	supervision	supervision	Basel	management	risk	Basel
regulation	regulation	risk	banking	banking	financial	supervision	supervision	supervision	supervision
risk	risk	regulation	regulation	financial	banking	management	Basel	management	management
financial	financial	financial	financial	regulation	management	banking	banking	regulation	regulation
loan	institution	institution	institution	institution	regulation	regulation	regulation	banking	banking
institution	loan	loan	management	management	community	organization	organization	institution	institution
community	system	small_business	community	community	institution	community	institution	loan	community
small_business	community	system	loan	credit	credit	financial	credit	community	organization
Words_2007	**Words_2008**	**Words_2009**	**Words_2010**	**Words_2011**	**Words_2012**	**Words_2013**	**Words_2014**	**Words_2015**	**Words_2016**
bank	bank	bank	bank	bank	bank	bank	bank	bank	bank
risk	risk	regulation	regulation	regulation	regulation	regulation	regulation	regulation	regulation
supervision	institution	supervision	supervision	community	community	supervision	supervision	supervision	supervision
regulation	supervision	institution	community	supervision	supervision	community	community	financial	firm
Basel	regulation	loan	loan	financial	firm	firm	firm	community	requirement
institution	loan	community	institution	firm	financial	requirement	requirement	liquidity	financial
loan	mortgage	financial	firm	institution	banking	financial	financial	firm	liquidity
banking	credit	firm	financial	loan	institution	institution	risk	requirement	capital
credit	liquidity	risk	banking	capital	requirement	capital	liquidity	institution	community
management	lender	lender	lender	banking	systemic	banking	banking	banking	stress_test

Table 9 – Interpretation of the three topics of the estimated LDA model from 1997 until 2016. See Appendix tables A1-A20 for the top 10 wordlists per topic per year.

Year	Interpretation topic 1	Interpretation topic 2	Interpretation topic 3
1997	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
1998	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
1999	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2000	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2001	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2002	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2003	Internal control, corporate governance and compliance	Monetary policy	Risks, regulations, and supervision of the banking system
2004	(global) economic and business developments	Monetary policy	Internal control, corporate governance and compliance
2005	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2006	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2007	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2008	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2009	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2010	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2011	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2012	(global) economic and business developments	Monetary policy	Financial stability and systemic risk
2013	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2014	Financial stability and systemic risk	Monetary policy	Risks, regulations, and supervision of the banking system
2015	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system
2016	(global) economic and business developments	Monetary policy	Risks, regulations, and supervision of the banking system

APPENDIX: Word lists from LDA analysis

Table A1 – Top 10 words of topics in 1997

1997 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
economy	rate	bank
community	inflation	financial
system	economy	risk
development	price	regulation
service	policy	banking
new	growth	supervision
loan	monetary_policy	activity
work	increase	system
Federal_Reserve	interest_rate	management
need	change	institution
1997 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
community	inflation	bank
system	rate	regulation
loan	monetary_policy	banking
service	decline	financial
payment	price	risk
development	forecast	supervision
affordable_housing	growth	activity
lender	unemployment	small_business
CRA	interest_rate	system
education	trend	institution

Table A2 – Top 10 words of topics in 1998

1998 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
bank	economy	bank
new	inflation	risk
service	rate	financial
loan	growth	capital
Federal_Reserve	price	supervision
may	policy	banking
business	monetary_policy	system
payment	change	regulation
system	forecast	international
financial	increase	country
1998 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
payment	inflation	supervision
service	monetary_policy	bank
CRA	rate	banking
small_business	forecast	international
community	growth	capital
technology	price	risk
lender	target	financial
consumer	decline	regulation
income	meeting	crisis
mortgage	interest_rate	standard

Table A3 – Top 10 words of topics in 1999

1999 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
economy	rate	bank
technology	economy	financial
new	inflation	risk
work	price	banking
business	policy	institution
international	monetary_policy	system
world	growth	supervision
trade	risk	regulation
good	increase	capital
country	interest_rate	credit
1999 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
technology	monetary_policy	bank
trade	rate	supervision
people	inflation	risk
economy	price	banking
farm	interest_rate	financial
job	target	small_business
social_security	growth	regulation
euro	risk	credit
american	forecast	institution
save	equity	CRA

Table A4 – Top 10 words of topics in 2000

2000 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
economy	bank	risk
rate	inflation	bank
growth	Federal_Reserve	financial
increase	central	supervision
technology	financial	management
investment	community	institution
new	target	system
high	system	capital
inflation	policy	banking
business	monetary_policy	regulation
2000 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
productivity	target	risk
economy	bank	bank
productivity_growth	inflation	supervision
inflation	payment	management
rise	central	institution
growth	community	capital
capital	monetary_policy	banking
demand	independence	regulation
supply	social_security	financial
rate	consumer	crisis

Table A5 – Top 10 words of topics in 2001

2001 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
economy	financial	bank
financial	money	risk
capital	rate	inflation
consolidation	asset	capital
growth	price	policy
increase	could	supervision
firm	debt	objective
rate	system	central
effect	use	target
investment	new	banking
2001 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
consolidation	money	bank
productivity	debt	risk
productivity_growth	treasury	supervision
study	reserve	objective
spending	currency	inflation
high_tech	social_security	Basel
firm	loan	accord
stock	mortgage	capital
slowdown	income	approach
service	private	practice

Table A6 – Top 10 words of topics in 2002

2002 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
bank	economy	financial
risk	policy	risk
financial	rate	bank
banking	price	regulation
management	inflation	credit
community	growth	institution
credit	monetary_policy	system
new	central	loan
business	rule	change
economy	stock	firm
2002 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
community	inflation	financial
banking	policy	regulation
bank	monetary_policy	credit
small_business	output	risk
credit	price	institution
financial	productivity	derivative
risk	rate	option
institution	spending	mortgage
supervision	economy	earnings
education	uncertainty	shareholder

Table A7 – Top 10 words of topics in 2003

2003 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
risk	inflation	bank
management	economy	risk
bank	price	financial
process	rate	capital
company	investment	Basel
internal_control	policy	credit
control	increase	banking
financial	low	supervision
report	growth	regulation
business	monetary_policy	system
2003 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
internal_control	inflation	Basel
company	price	supervision
control	rate	bank
corporate_governance	monetary_policy	banking
management	stock	capital
japan	target	regulation
process	stability	consumer
risk	central	education
report	interest_rate	foreign
director	economy	credit

Table A8 – Top 10 words of topics in 2004

2004 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
economy	rate	risk
united_state	economy	bank
country	price	financial
debt	inflation	management
increase	policy	credit
trade	monetary_policy	business
job	increase	firm
rise	low	institution
capital	central	new
rate	interest_rate	organization
2004 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
country	inflation	risk
job	monetary_policy	management
trade	rate	credit
deficit	oil	bank
debt	price	organization
current_account	policy	customer
economy	central	corporate_governance
productivity	FOMC	internal_control
currency	expectation	compliance
foreign	economy	transaction

Table A9 – Top 10 words of topics in 2005

2005 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
investment	price	bank
economy	economy	risk
rate	inflation	Basel
price	rate	institution
increase	policy	capital
asset	increase	regulation
current_account	central	financial
deficit	financial	system
low	growth	banking
rise	energy	credit
2005 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
current_account	inflation	bank
deficit	energy	Basel
rate	rate	risk
investment	economy	banking
rise	price	supervision
save	monetary_policy	institution
productivity_growth	monetary	management
home	recession	regulation
dollar	target	agency
mortgage	oil	credit

Table A10 – Top 10 words of topics in 2006

2006 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
financial	economy	risk
bank	price	bank
economy	inflation	management
community	rate	capital
development	increase	Basel
system	growth	supervision
Federal_Reserve	high	regulation
investment	recent	banking
work	low	institution
information	investment	organization
2006 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
community	inflation	risk
check	price	management
economy	rate	Basel
financial	economy	compliance
reserve	energy	supervision
bank	productivity	regulation
crisis	demand	framework
cash	growth	banking
private	labor	bank
development	long_term	banker

Table A11 – Top 10 words of topics in 2007

2007 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
financial	inflation	bank
risk	economy	risk
investment	rate	financial
economy	price	regulation
credit	monetary_policy	institution
liquidity	bank	lender
increase	central	system
capital	expectation	Federal_Reserve
country	policy	consumer
asset	effect	banking
2007 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
liquidity	inflation	supervision
risk	monetary_policy	regulation
investment	price	Basel
financial	expectation	community
flow	rate	banking
price	housing	bank
current_account	house_price	payment
credit	output	agency
premium	long_run	disclosure
emerge	employment	risk

Table A12 – Top 10 words of topics in 2008

2008 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
financial	inflation	risk
credit	price	mortgage
economy	economy	loan
bank	rate	bank
liquidity	monetary_policy	management
risk	policy	consumer
Federal_Reserve	bank	institution
institution	objective	borrower
asset	central	lender
fund	expectation	regulation
2008 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
financial	inflation	mortgage
credit	monetary_policy	management
liquidity	price	loan
economy	expectation	risk
fund	economy	foreclosure
turmoil	objective	community
loss	policy	consumer
security	bubble	borrower
funding	rate	credit
institution	output	supervision

Table A13 – Top 10 words of topics in 2009

2009 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
economy	financial	financial
financial	Federal_Reserve	bank
credit	credit	firm
mortgage	bank	risk
policy	economy	regulation
loan	asset	supervision
consumer	lender	community
growth	rate	capital
price	program	institution
lender	institution	system
2009 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
economy	central	community
consumer	Federal_Reserve	supervision
recession	Fed	foreclosure
quarter	balance_sheet	regulation
panic	short_term	firm
growth	treasury	requirement
trade	credit	organization
intermediation	interest_rate	CRA
product	program	local
inflation	facility	big_fail

Table A14 – Top 10 words of topics in 2010

2010 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
credit	economy	financial
bank	policy	bank
loan	rate	firm
small_business	inflation	risk
community	monetary_policy	system
consumer	fiscal	crisis
lender	growth	regulation
economy	financial	economy
Federal_Reserve	country	capital
business	increase	institution
2010 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
small_business	inflation	firm
community	fiscal	financial
loan	monetary_policy	risk
credit	emerge	liquidity
consumer	policy	central
business	foreclosure	bank
credit_card	global	institution
banker	price	funding
meeting	long_term	regulation
lender	house_price	capital

Table A15 – Top 10 words of topics in 2011

2011 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
community	economy	financial
loan	rate	bank
credit	policy	capital
business	inflation	firm
mortgage	increase	risk
small_business	financial	regulation
bank	price	system
housing	growth	institution
consumer	long_term	economy
Federal_Reserve	recovery	crisis
2011 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
loan	inflation	capital
community	rate	regulation
small_business	fiscal	financial
family	long_term	requirement
mortgage	recovery	supervision
foreclosure	purchase	systemic
housing	FOMC	firm
wealth	economy	liquidity
property	unemployment	system
income	monetary_policy	bank

Table A16 – Top 10 words of topics in 2012

2012 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
mortgage	economy	bank
housing	rate	financial
home	policy	regulation
credit	unemployment	community
lender	inflation	firm
economy	monetary_policy	risk
low	FOMC	system
recovery	long_term	capital
Federal_Reserve	federal	supervision
loan	Federal_Reserve	banking
2012 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
housing	inflation	bank
mortgage	rate	regulation
home	unemployment	community
lender	FOMC	supervision
recovery	committee	foreign
foreclosure	monetary_policy	banking
loan	policy	firm
property	economy	institution
household	labor	financial
credit	objective	stability

Table A17 – Top 10 words of topics in 2013

2013 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
bank	rate	financial
financial	economy	regulation
community	policy	firm
risk	purchase	capital
mortgage	monetary_policy	bank
loan	Federal_Reserve	risk
Federal_Reserve	committee	asset
lender	FOMC	large
credit	inflation	requirement
crisis	federal	liquidity
2013 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
community	rate	firm
bank	monetary_policy	regulation
loan	committee	liquidity
mortgage	FOMC	requirement
income	inflation	institution
lender	purchase	capital
household	interest_rate	funding
local	unemployment	systemic
network	long_term	fire_sale
regulation	economy	dealer

Table A18 – Top 10 words of topics in 2014

2014 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
financial	economy	bank
risk	policy	financial
monetary_policy	rate	regulation
stability	labor	firm
Federal_Reserve	recovery	risk
economy	inflation	supervision
rate	committee	capital
policy	growth	asset
household	job	community
low	continue	liquidity
2014 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
monetary_policy	labor	bank
family	rate	regulation
wealth	recovery	supervision
household	committee	liquidity
premium	FOMC	firm
rate	inflation	requirement
risk	purchase	capital
financial	job	community
credit	economy	risk
stability	growth	banking

Table A19 – Top 10 words of topics in 2015

2015 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
economy	inflation	bank
policy	rate	financial
growth	economy	regulation
crisis	price	risk
financial	federal	system
monetary_policy	low	institution
rate	fund	firm
global	monetary_policy	liquidity
central	level	large
country	FOMC	capital
2015 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
growth	inflation	regulation
monetary_policy	rate	liquidity
rate	federal	bank
economy	FOMC	institution
exchange	price	system
china	labor	supervision
gdp	economy	firm
world	long_run	requirement
global	expectation	banking
country	monetary_policy	lender

Table A20 – Top 10 words of topics in 2016

2016 – Word probabilities top 10		
Topic 1	**Topic 2**	**Topic 3**
economy	rate	financial
inflation	economy	bank
growth	interest_rate	capital
rate	monetary_policy	firm
low	policy	regulation
policy	fund	risk
decline	federal	liquidity
increase	financial	requirement
work	price	system
labor	technology	asset
2016 – Term scores top 10		
Topic 1	**Topic 2**	**Topic 3**
growth	rate	regulation
inflation	interest_rate	liquidity
rate	monetary_policy	requirement
economy	federal	firm
low	interest	capital
labor	equilibrium	stress_test
job	committee	bank
china	long_run	system
oil	FOMC	financial
labor_force	economy	stress

SUPPLEMENTAL ONLINE APPENDIX

What Say They About Their Mandate? A Textual Assessment of Federal Reserve Speeches

Myrthe van Dieijen and Robin L. Lumsdaine

ONLINE APPENDIX A: Details of Data Cleaning

Table OA1 - Steps of the data cleaning process with the respective number of tokens and types in the texts after the step has been performed.

Data cleaning steps	Number of tokens	Number of types
Lower cased texts that are cleared from punctuation and (numerical) symbols	3,788,490	24,116
Lemmatize all words[37]	-0	-5885 + 997 = -4888
Remove stopwords[38]	-1,643,201	-149
Remove words with a length below 3 characters	-15,918	-173
Create phrases (bigrams) [39]	-197,618	-1,188 + 20,405 = +19,217
Prune the texts from words that occur with a high frequency (i.e., in more than 95 percent of the speeches)	-82,492	-9
Cleaned texts	**1,849,261**	**38,114**

[37] After lemmatization 5,885 unique words (types) cannot be found in the texts anymore, as they are brought back to their basic form (e.g., 'doing' becomes 'do'). Lemmatization creates 997 new unique words as well, as some words only occurred in the texts in an inflected form.

[38] The total number of stopwords that was searched for in the texts is 165 (see Table OA2 for the detailed list). Some stopwords we searched for were not found in the speeches, as those were already removed from the texts due to the lemmatization procedure, those words were: 'mustn', 'are', 'has', 'its', 'shan', 'had', 'were', 'is', 'having', 'does', 'been', 'mightn', 'did', 'was', 'am', 'doing'. Hence, we removed 149 of the 165 stopwords from the texts.

[39] Most of the bigrams that were created (20,405) consist of words that occur both as part of the phrase as well as individually. For example, both the words 'productivity' and 'growth', as well as the bigram 'productivity_growth' can be found in the texts after the bigrams were created. However, some words (1,188 to be precise) are no longer found in the texts individually after the bigram creation procedure. For example, 'york' does not occur as a separate word following the bigram creation procedure because everywhere it occurred in a text it was alongside the word 'new' (as in 'new york'), and it is therefore now enveloped into the phrase 'new_york'. The word `new' still appears as an individual word as well as in the bigram `new_york'.

Table OA2 – Stopwords list

Stopwords					
NLTK version 3.2.1 stopwords					
a	can	here	myself	shouldn	was
about	couldn	hers	needn	so	we
above	d	herself	no	some	were
after	did	him	nor	such	weren
again	didn	himself	not	t	what
against	do	his	now	than	when
ain	does	how	o	that	where
all	doesn	i	of	the	which
am	doing	if	off	their	while
an	don	in	on	theirs	who
and	down	into	once	them	whom
any	during	is	only	themselves	why
are	each	isn	or	then	will
aren	few	it	other	there	with
as	for	its	our	these	won
at	from	itself	ours	they	wouldn
be	further	just	ourselves	this	y
because	had	ll	out	those	you
been	hadn	m	over	through	your
before	has	ma	own	to	yours
being	hasn	me	re	too	yourself
below	have	mightn	s	under	yourselves
between	haven	more	same	until	
both	having	most	shan	up	
but	he	mustn	she	ve	
by	her	my	should	very	
Additional stopwords used in this paper					
percent	first	third	fifth	seventh	ninth
number	second	fourth	sixth	eighth	tenth

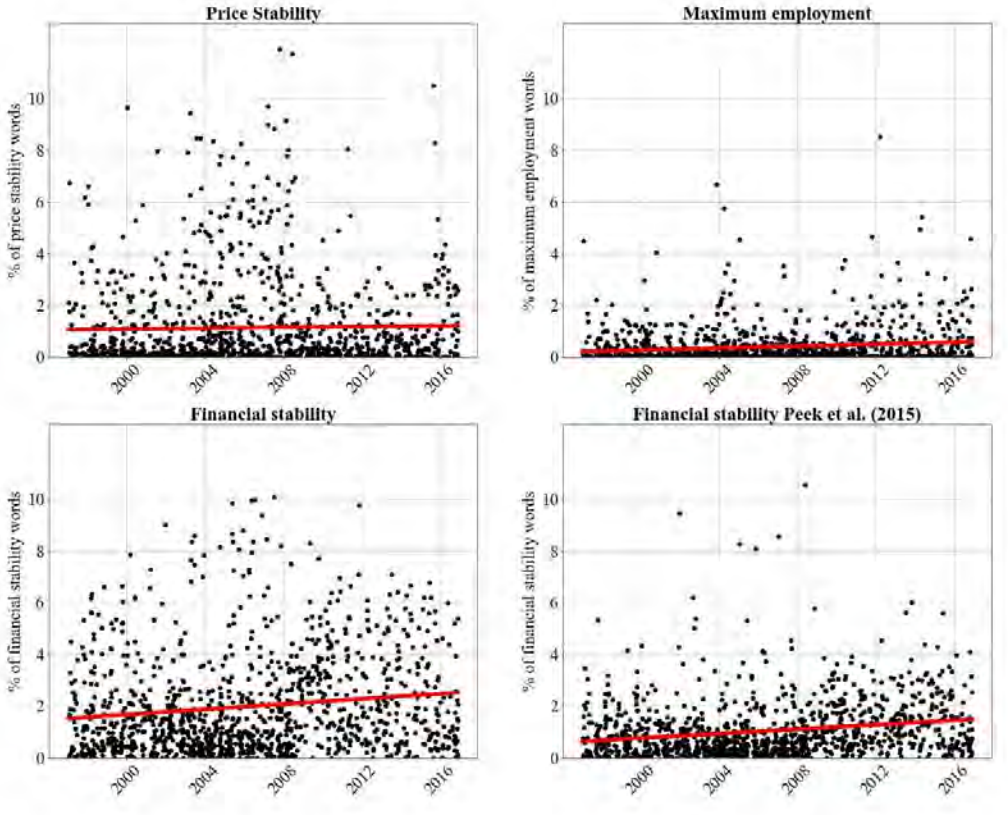
Table OA3 – Pruned words that occurred in more than 95 percent of the speeches

Pruned high frequency words	
also	well
make	would
many	one
market	year
time	

Table OA4 – Grouping of related words

Word combinations	
banking/banking_organization	investment/investor
borrower/borrow/borrowing	lender/lend/lending
consumer/consumption	management/manager
development/developmental	organization/organizational
disinflation/disinflationary	policy /policymaker/policy_maker
economy/economic/economics	regulation/regulator/regulatory
employment/employed	risk/risky/riskiness
financial/financially	supervision/supervisor/supervisory
government/governmental	technology/technological/technologically
inflation/inflationary	unemployment/unemployed
institution/institutional	work/worker

Figure OA1 – Percentage of words of the price stability, maximum employment, and two types of financial stability dictionary categories per speech over time. The line represents the least squares polynomial fit.



ONLINE APPENDIX B: Variational Inference

LDA estimation through Variational Inference

The key inferential problem of LDA is to compute the posterior distribution of the latent variables β, θ and $\boldsymbol{z}_{1:D}$ conditional on the observed corpus $\boldsymbol{w}_{1:D}$, α and η:

$$p(\theta, \beta, \boldsymbol{z}_{1:D}|\boldsymbol{w}_{1:D}, \alpha, \eta) = \frac{p(\theta, \beta, \boldsymbol{z}_{1:D}, \boldsymbol{w}_{1:D}|\alpha, \eta)}{p(\boldsymbol{w}_{1:D}|\alpha, \eta)} = \frac{p(\theta|\alpha)p(\boldsymbol{w}_{1:D}|\boldsymbol{z}_{1:D}, \beta)p(\boldsymbol{z}_{1:D}|\theta)}{p(\boldsymbol{w}_d|\alpha, \beta)} \quad (1)$$

The joint distribution in the numerator can be written as:

$$p(\theta, \beta, \boldsymbol{z}_{1:D}, \boldsymbol{w}_{1:D}|\alpha, \eta) = \left(\prod_{k=1}^{K} p(\beta_k|\eta)\right)\left(\prod_{d=1}^{D} p(\theta_d|\alpha)\prod_{n=1}^{N_d} p(z_{d,n}|\theta_d)p(w_{d,n}|\beta, z_{d,n})\right) \quad (2)$$

The denominator, which is known as the *evidence*, can be obtained by marginalizing over the latent variables β, θ and $\boldsymbol{z}_{1:D}$:

$$p(\boldsymbol{w}_{1:D}|\alpha, \eta) = \int\int \sum_{\boldsymbol{z}_{1:D}} \left(\prod_{k=1}^{K} p(\beta_k|\eta)\right)\left(\prod_{d=1}^{D} p(\theta_d|\alpha)\prod_{n=1}^{N_d} p(z_{d,n}|\theta_d)p(w_{d,n}|\beta, z_{d,n})\right) d\theta d\beta \quad (3)$$

The numerator can be computed easily, but the *evidence* is intractable to compute as the latent variables β and θ are connected via the observed corpus (Blei, 2011).[40] Consequently, the posterior in (1) cannot be analytically computed, but there are algorithms that can be used to approximate it. Topic modelling algorithms are typically either sampling-based (e.g., Gibbs sampling) or variational.[41] There has been much debate about which of the two approaches is better, and it depends – among other things – on the model being used (see Asuncion et al., 2009 for an overview of the pros and cons of both methods). In this paper, we use a variational algorithm, because it is more suitable for the dynamic estimation of the DTM than sampling methods (Blei et al., 2006).

[40] Using the definitions and terminology of section 4.1 and 4.2, the evidence can be written as: $p(\boldsymbol{w}_{1:D}|\alpha, \eta) = \int\int \left(\prod_{k=1}^{K} \frac{\Gamma(\sum_{v=1}^{V}\eta_v)}{\prod_{v=1}^{V}\Gamma(\eta_v)} \prod_{v=1}^{V} \beta_{k,v}^{\eta_v - 1}\right) \prod_{d=1}^{D} \left(\frac{\Gamma(\sum_{k=1}^{K}\alpha_k)}{\prod_{k=1}^{K}\Gamma(\alpha_k)} \prod_{k=1}^{K} \theta_{d,k}^{\alpha_k - 1}\right) \left(\prod_{n=1}^{N_d} \sum_{k=1}^{K} \prod_{v=1}^{V} (\theta_{d,k}\beta_{k,v})^{w_{d,n}^{v}}\right) d\theta d\beta$. This probability can be obtained by maximizing the log likelihood of the observed corpus $\boldsymbol{w}_{1:D}$ given the fixed values for α and η. However, in computing the logarithm of this function $\theta_{d,k}$ and $\beta_{k,v}$ are not separable in summing over all the possible values for the latent topic structure. As there are an exponentially large number of possible topic structures, this sum is intractable to compute (Blei, 2011).

[41] Sampling-based algorithms aim to collect samples from the intractable posterior by approximating it with an empirical distribution. Variational methods posit a parameterized family of distributions over the latent topic structure and then find the member of that family that is closest to the posterior (Blei et al., 2011).

The main goal of variational inference is to find a simpler, tractable distribution that obtains a *lower bound* on the log likelihood of the corpus ($\log p(\boldsymbol{w}_{1:D}|\alpha,\eta)$), a bound that can be maximized with respect to α and η via the Expectation Maximization (EM) algorithm. To obtain the tightest possible lower bound, a tractable family of variational distributions on the latent variables is required. This family of distributions together makes up the following mean field variational distribution for LDA:

$$q(\theta,\beta,\boldsymbol{z}_{1:D}) = \prod_{k=1}^{K} q(\beta_k|\lambda_k) \prod_{d=1}^{D} q(\theta_d|\gamma_d) \prod_{n=1}^{N_d} q\left(z_{d,n}|\phi_{d,n}\right) \quad (4)$$

Where λ_k, γ_d and $\phi_{d,n}$ are the free *variational parameters* of the variables β_k, θ_d and $z_{d,n}$. In contrast to the true posterior, the latent variables are not coupled anymore, but are now independent of one another.

After specifying a tractable variational distribution the variational parameters can be obtained by minimizing the Kullback-Leibler (KL) distance to the true posterior:

$$\underset{\gamma,\lambda,\boldsymbol{\phi}_{1:D}}{\operatorname{argmin}} \; KL\left(q(\theta,\beta,\boldsymbol{z}_{1:D})||p(\theta,\beta,\boldsymbol{z}_{1:D}|\boldsymbol{w}_{1:D})\right) \quad (5)$$

Minimizing this distance is equivalent to maximizing the lower bound on the log likelihood. This lower bound is called the *evidence lower bound* (ELBO), and it's defined as (Srivastava and Sahami, 2009):

$$\mathcal{L} = \sum_{k=1}^{K} E[\log p(\beta_k|\eta)] + \sum_{d=1}^{D} E[\log p(\theta_d|\alpha)] + \sum_{d=1}^{D}\sum_{n=1}^{N_d} E\left[\log p\left(z_{d,n}|\theta_d\right)\right] + \sum_{d=1}^{D}\sum_{n=1}^{N_d} E\left[\log p\left(w_{d,n}|\beta,z_{d,n}\right)\right] + H(q) \quad (6)$$

This objective cannot be computed exactly, but only up to the constant $H(q)$, which is independent of the variational parameters.[42] Optimizing this lower bound proceeds by coordinate ascent. Before starting the coordinate ascent updating procedure the values for K, α and η are set to the values $K = 3$, $\alpha = 50/K$, and $\eta = 0.01$ (as described in section 4.3). For details on how the variational parameters are updated in each iteration of the mean field variational inference algorithm, see Srivastava and Sahami (2009).

[42] The constant is equal to the log likelihood of the observed corpus (see Mandt and Blei, 2014).

DTM estimation through Variational Inference

As with LDA, the posterior distribution of the DTM is intractable to compute, and has to be approximated through variational inference.[43]

As noted above, with variational inference the aim is to find a distribution that is as close as possible to the true posterior by minimizing the Kullback-Liebler (KL) divergence. In DTM we have three latent variables: the topics $\beta_{t,k}$, the document-topic proportions $\theta_{t,d}$ and the topic-assignments $z_{t,d,n}$. Consequently, there are a variational parameters for each of these variables. Because the latent topics $\beta_{t,k}$ evolve over time, for each topic's sequence of multinomial parameters there is a variational parameter. The approximate variational posterior is:

$$\prod_{k=1}^{K} q(\beta_{k,1}, \ldots, \beta_{k,T} | \hat{\beta}_{k,1}, \ldots, \hat{\beta}_{k,T}) \times \prod_{t=1}^{T} \left(\prod_{d=1}^{D} q(\theta_{t,d} | \gamma_{t,d}) \prod_{n=1}^{N_{t,d}} q(z_{t,d,n} | \phi_{t,d,n}) \right)$$

This distribution is Gaussian, and the parameters are fit to minimize the KL divergence compared to the true (not necessarily Gaussian) posterior. The time-varying structure of each topic is incorporated in the variational distribution of $\{\beta_{k,1}, \ldots, \beta_{k,T}\}$ (Blei et al., 2006). The variational distributions of the document-level latent variables are exactly the same as in the LDA model (Blei et al., 2003). The variational approximation of the topic parameters in the DTM can be done via a Kalman filter.[44] To explain Variational Kalman Filtering in a simple way, Blei et al. (2006) explained the method for a unigram model, where β_t is the only latent parameter. In that scenario, they approximate the posterior $p(\beta_{1:T} | \boldsymbol{w}_{1:D,1:T})$ using the state space posterior $q(\beta_{1:T} | \hat{\beta}_{1:T})$. For more details on the Variational Kalman Filtering algorithm, see Blei et al. (2006).

[43] Due to the coupling between β and θ through the observed documents, the posterior (i.e., the evidence of the posterior) is intractable. Whereas sampling methods (e.g., Gibbs sampling) have proven to be effective for LDA, the nonconjugacy of the Gaussian and multinomial distributions in DTM renders such methods inviable, thus necessitating use of variational inference.

[44] In this paper the Kalman filter is used, as that is the approximation method in the DTM model of the gensim Python package, version 0.13.3. Alternatively, the approximation can be done via a wavelet regression (Blei et al., 2006).

ONLINE APPENDIX C: Robustness checks document-topic proportions

Figure OC1 – Evolution of the topic-word probabilities of the words that occurred at least once in the top 5 of each of the three topics from 1997 to 2016.

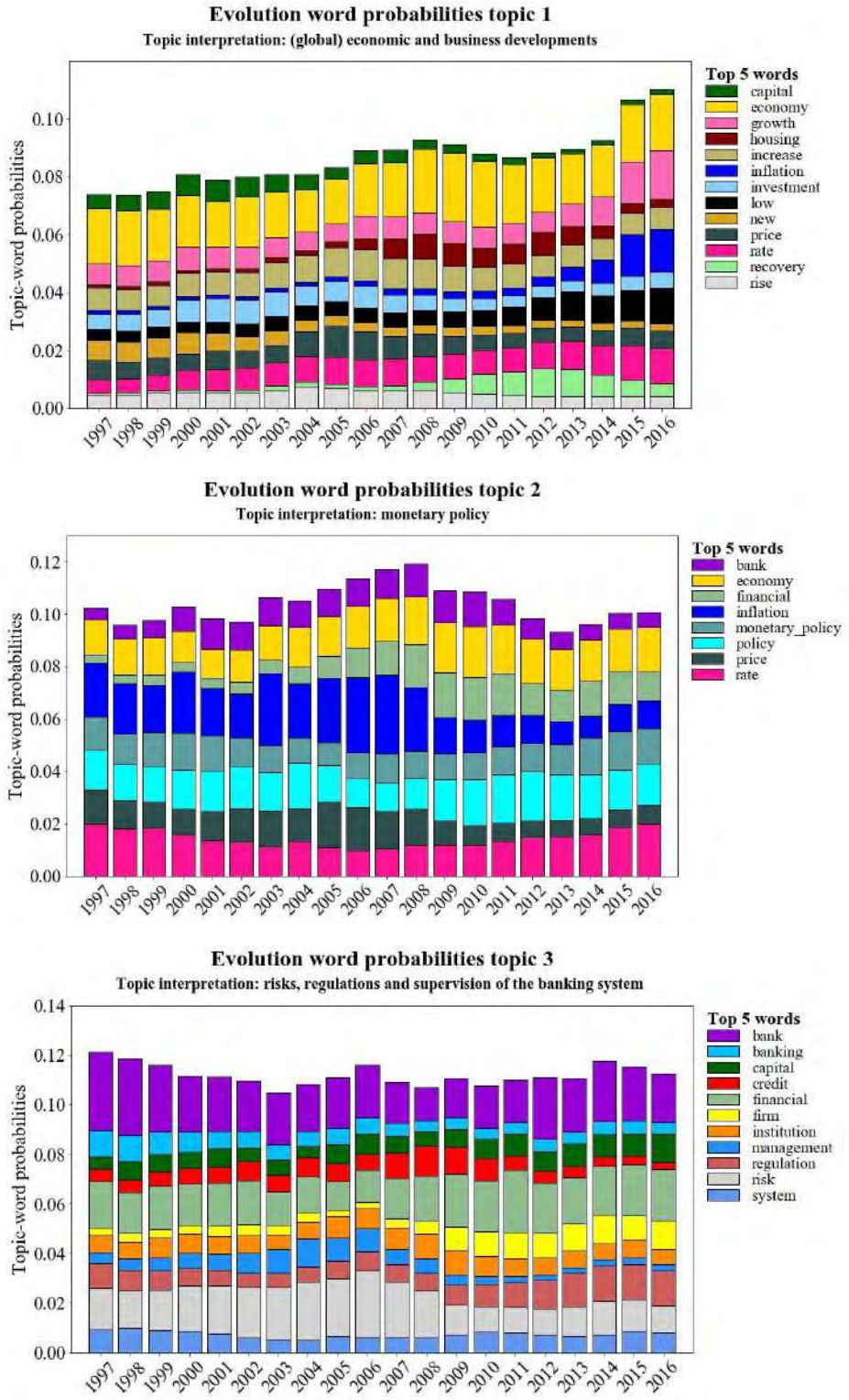


Figure OC2 –The average document-topic proportions per speech from 1997 to 2016 (before computing the average, document-topic proportions below 0.05 were omitted and the remaining proportions per speech were normalized to sum to 1), as estimated by the DTM.

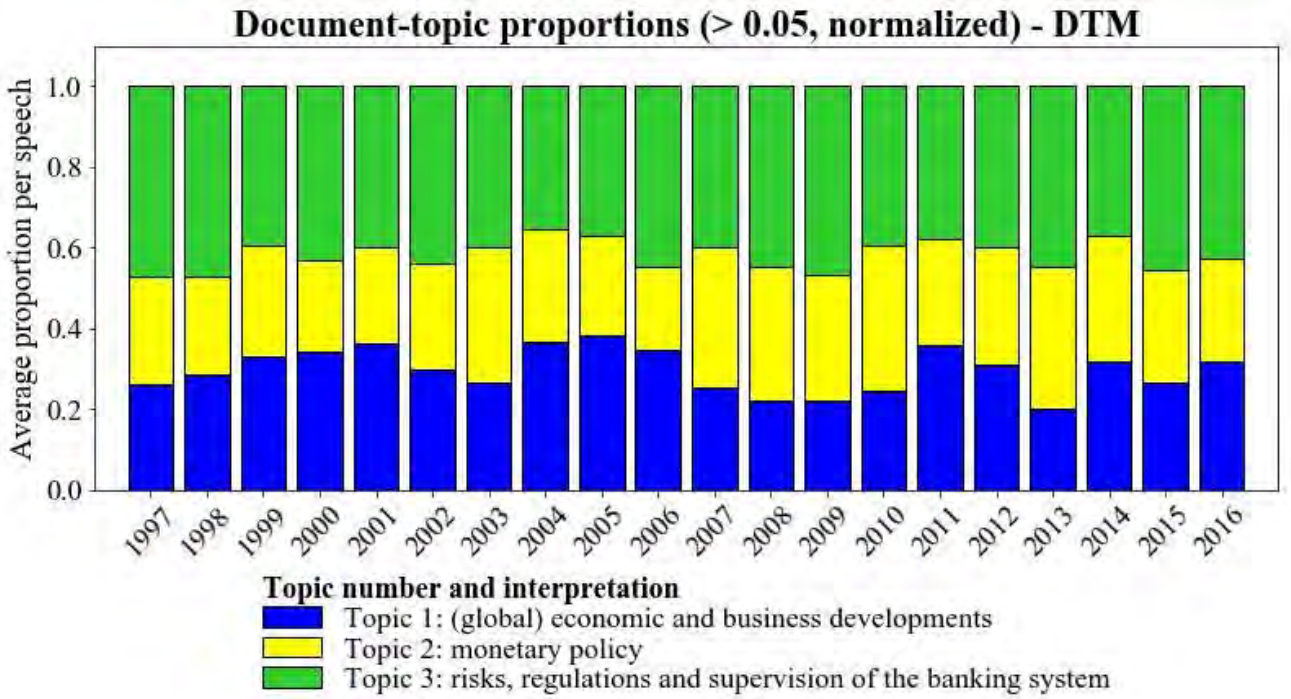


Figure OC3–The percentage of speeches devoted to each topic, classified according to which topic had the highest document-topic proportion for that speech, by year, as estimated by the DTM.

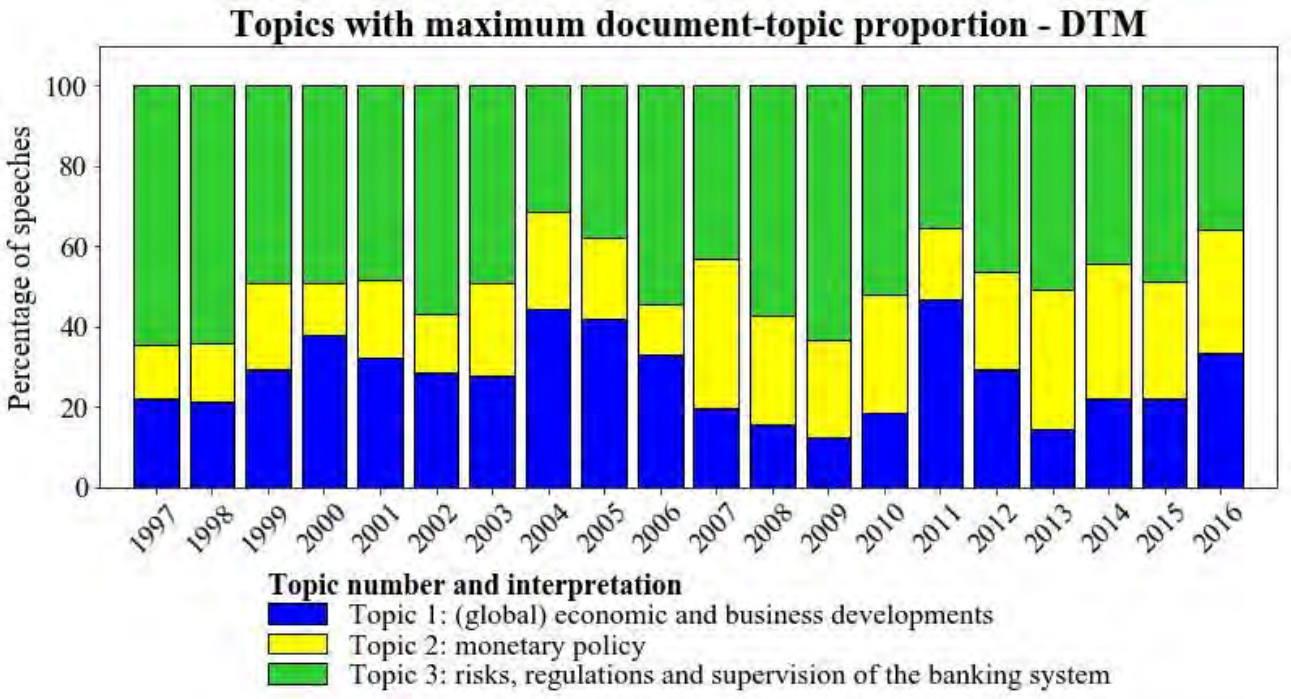


Figure OC4 – The average document-topic proportions per speech from 1997 to 2016 (where for each topic, speeches with document-topic proportions below 0.05 were omitted from the calculation of the yearly average, after which the average proportions per topic per year were normalized to sum to 1), as estimated by the LDA model.

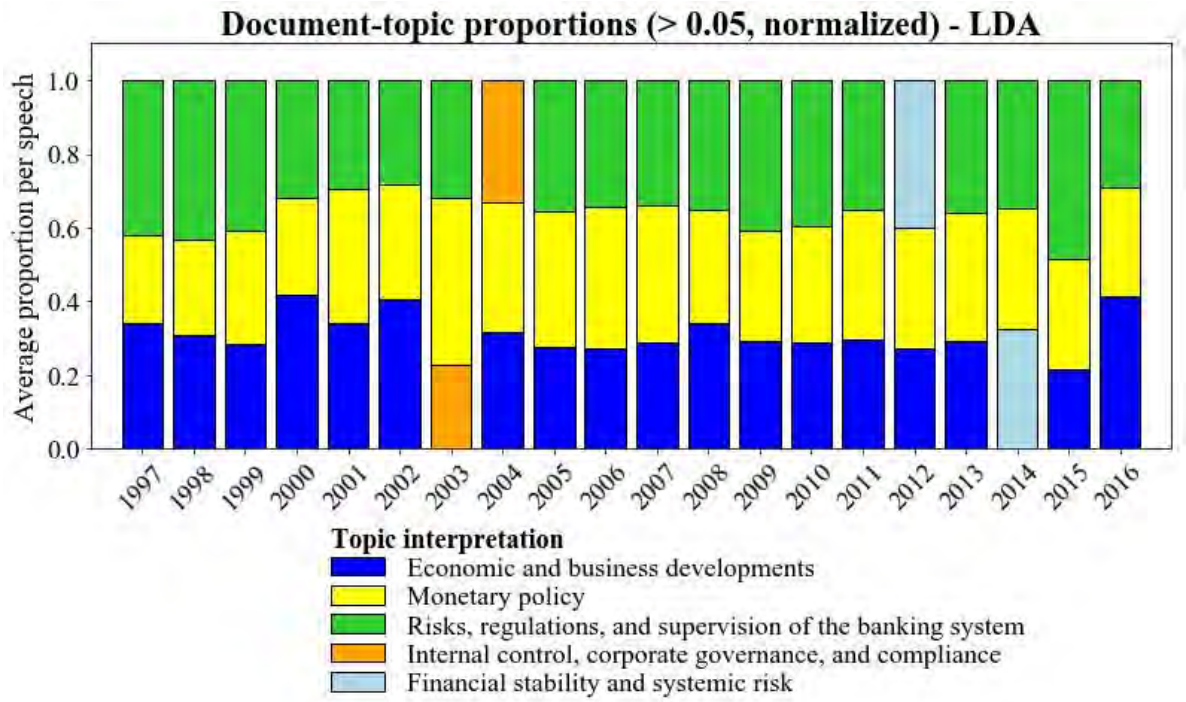
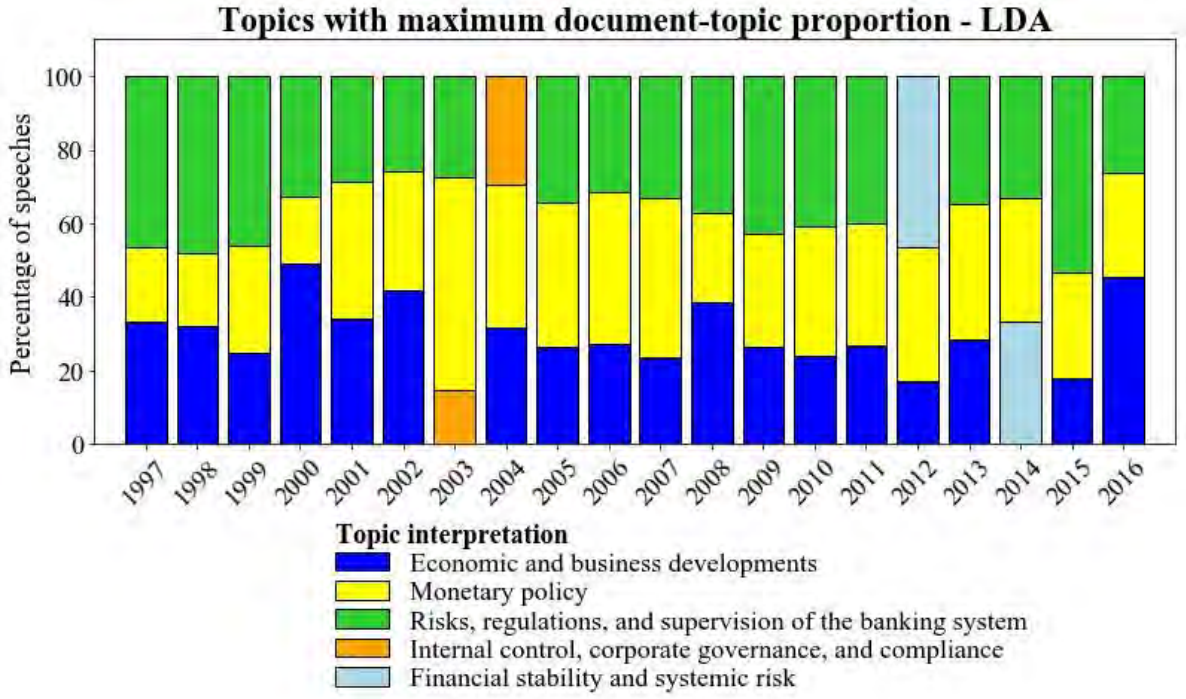


Figure OC5 – Percentage of speeches devoted to each topic, classified according to which topic had the highest document-topic proportions for that speech, by year, as estimated by the LDA model.



ONLINE APPENDIX D: Word clouds showing topic assignments

Figure OD1 – The word clouds of the top 100 words (based on highest word probabilities) of the three topics before, during, and after the crisis, as estimated by the DTM. The size of each word is indicative of the probability and the color shows the topic in which it obtained the highest probability. The topic number is appended using an underscore to words that occurred in the top 100 of multiple topics. The left column contains word clouds based on the word probabilities and the right column contains the associated word clouds based on the termscores.