Deciphering Fedspeak: The Information Content of FOMC Meetings

Narasimhan Jegadeesh
Di (Andrew) Wu*

March 23, 2017

ABSTRACT

We present a new approach to quantify the economic and policy content of the Federal Reserve communications by dissecting the Federal Open Market Committee (FOMC) meeting minutes into eight distinct economic topics. We examine the informativeness of the Fed's discussion of each of these topics for the stock market and for interest rates. The market finds the Fed's discussion of its policy stance, inflation and employment to be the most informative and its discussion of topics such as trade, consumption and investment are not informative.

JEL classification: G18, G28, E58

*Narasimhan Jegadeesh is the Dean's Distinguished Professor at the Goizueta Business School, Emory University, ph: 404-727-4821, email: jegadeesh@emory.edu. Di (Andrew) Wu is Assistant Professor of Technology, Operations, and Finance at the Stephen M. Ross School of Business, University of Michigan, ph: 734-765-4178, email: andydiwu@umich.edu. We would like to thank Itay Goldestein, Gerard Hoberg, Tim Loughran, David Lucca, Bill McDonald, Jeremy Siegel, Rob Stambaugh, Paul Tetlock, Paula Tkac, Amir Yaron, and seminar participants at the Federal Reserve Bank of Atlanta and the Wharton School for helpful comments and suggestions.

Monetary policies implemented by the Federal Reserve (Fed) have a significant impact on both the real economy and on the prices of financial assets, including stocks and bonds. While real macroeconomic variables, such as GDP and employment, respond to the effects of policy innovations over a long horizon, the financial market rationally anticipates future changes to the real economy and financial asset prices react instantaneously. The Federal Open Market Committee (FOMC), which is comprised of twelve members, is responsible for the open market operations that the Fed conducts to implement its monetary policy. The FOMC meets eight times a year and at its meetings, "the Committee reviews economic and financial conditions, determines the appropriate stance of monetary policy, and assesses the risks to its long-run goals of price stability and sustainable economic growth."[1]

The FOMC releases the minutes of its meetings about three weeks after the meetings. The market learns about FOMC's assessment of the economy and its monetary policy stance through these minutes. This paper applies a new approach to analyze the textual content of the minutes and examines its information content for financial markets.

The Fed discusses a variety of topics during its meetings. For example, the Fed may discuss topics such as inflation, interest rates, employment and trade. Different topics likely have different levels of informativeness and understanding these differences will help us assess the importance that the market attaches to various topics. One of the important goals of this paper is to algorithmically identify the topics that the FOMC discusses in its meetings, and determine their relative importance for the market. We use the Latent Dirichlet Allocation (LDA) model of automated topic retrieval for topic identification. The LDA has been successfully employed to characterize topics from a wide variety of document sources, including journal articles in *Nature* to patient-discharge reports. We use this algorithm to identify the topics or mixture of topics in each paragraph of the minutes.

We identify the following eight topics in the minutes: policy stance, inflation, financial market, employment, economic growth, foreign trade, consumption and production and investment. The proportion of the minutes devoted to each of these topics and the tone vary across meetings. The tone of the FOMC's minutes is generally negative on all topics during recessions, and it is also negatively correlated to the level of unemployment.

We first examine how the stock and bond markets react to proportions of discussion devoted to each topic. We use SPY, an actively traded ETF that tracks the S&P500 Index, and LIBOR, implied by the

[1]Source: https://www.federalreserve.gov/monetarypolicy/fomc.htm

Eurodollar futures contract prices, as proxies for stock and bond markets, respectively. We find that variations in the proportion of the traditional "dual mandate" topics of inflation and employment, as well as the policy stance topic, are strongly related to the magnitude of price reactions immediately following the release of the minutes.

Next, we determine the tone of the minutes in its entirety and the tone of each individual topic based on the tonal words in the minutes. We find that when we consider the entire minutes as an individual unit, the tone of the minutes is only marginally informative. However, when we consider the tone of individual topics, we find that the contents of some topics are more informative than others. For example, discussions of policy stance and inflation are informative, but discussions of other topics, such as trade and consumption, are not. We also examine the directional impact of tone of each topic. We find that the tone of policy stance is positively correlated with changes in SPY but negatively correlated with changes in LIBOR. These results indicate that when the Fed takes up an easing posture, the stock market tends to go up and the interest rate tends to go down. We find similar results for inflation as well.

Our results indicate that the information content varies across topics. Also, in the case of stocks, the relation between the tone of some topics and direction of stock price changes are significantly positive, while the relation is negative for some other topics. These finer details about the incremental information content are not evident when we examine the document in its entirety, and these results illustrate the importance of topic level identification and analysis. Our results are robust within subperiods, and with respect to changes in the choice of tonal lexicons and in the number of topics.

This paper is one of the first in economics and finance to employ the Latent Dirichlet Allocation (LDA) model for automated paragraph-level topic identification. Our approach moves beyond the traditional bag of words approach employed in current literature, such as Tetlock (2007), Hanley and Hoberg (2010), Loughran and McDonald (2011), and Jegadeesh and Wu (2013). The LDA approach is particularly suitable for textual analysis of documents such as the FOMC minutes that contain a variety of topics, with significant variations in information content across topics.

Our paper also adds to the literature that characterizes Fed policy, and examines the relation between monetary policy innovations and the real economy. The literature follows a variety of different methods to examine the implication of Fed policy innovations. For example, Friedman and Schwartz (1963) and Romer and Romer (1989) identify large monetary shocks and examine their real effects using what

the latter calls a "narrative approach." These papers focus exclusively on large shocks, which occur relatively infrequently, and they track the long-term trajectory of the macroeconomy following these shocks.

Another strand of the literature uses a quantitative approach and statistically examines the relation between monetary policy variables and real macroeconomic variables. For example, Sims and Zha (2006) examine the multivariate relations between money supply, the Fed funds rate and real macroeconomic variables, such as GDP and unemployment under different monetary regimes, using a VAR model. Ang, Boivin, Dong, and Loo-Kung (2002), Primiceri (2005), Sargent, Williams, and Zha (2006), and Campbell, Pflueger, and Viceira (2015) also use statistical or structural models to examine the effects of monetary policy.

Our paper focuses on identifying monetary policy changes conveyed through the text of the minutes of FOMC meetings and examines their impact on financial assets. Earlier papers by Fleming and Piazzesi (2005) and Lucca and Moench (2015) also examine market price reactions to the Fed statements released on the day of FOMC meetings, and Rosa (2013) examines the response of treasury yields to the release of minutes. While these papers broadly examine whether FOMC committee meetings are intrinsically informative, our paper focuses on identifying the types of information contained in the minutes and the informativeness of different topics discussed in the meetings.

The rest of the paper is organized as follows. Section I describes our sample and data sources. Section II introduces our topic-based content analysis methodology. Section III reports the results of our empirical tests and explores the sources of informativeness of our measures. Section IV establishes the robustness of our results, and Section V provides a conclusion to our discussion.

I. Data

A. Introduction to FOMC Meetings and Minutes

This subsection provides a brief overview of the logistic details of FOMC meetings and the release of the meeting minutes. From the early 1980s, the FOMC holds eight regularly scheduled meetings per year, during which members discuss the economic outlook and formulate monetary policy. Any policy change decided at the meeting is implemented through open market operations. Prior to 1994, no public announcement about policy was made, and the market inferred any policy change through the

size and direction of the Fed's subsequent open market operations. Starting from January 1994, specific policy changes were made public in a short meeting statement released immediately after the meeting.

Moreover, detailed records of the discussions during each meeting are summarized in the form of meeting minutes and released to the public after a delay.[2] The minutes contain no new information received between the meeting date and the release date, and instead serve as an overview of the members' internal discussions on their economic outlook, as well as a nuanced explanation of the rationale for any policy change.

The meeting minutes follow a structured writing style. They are organized into four major sections. The first section outlines the administrative details of the meeting and reviews previous open market operations. The second section provides the staff's review and outlook of the economic and financial situation, prepared in advance of the meeting. The next two sections provide the bulk of the economic content. The third section details the FOMC members' discussion of the current economic and financial situation, as well as their own economic outlook and projections. The last section is mostly related to policy and discusses the rationale for current policy and outlook for future policies. We remove the first section prior to processing the documents since it is unlikely to contain any economically meaningful content.

B. The FOMC Minutes Sample

We download all FOMC meeting minutes between the February 1991 and June 2015 meetings from the FOMC's website. Some minutes in earlier periods are only available in scanned PDF format, and we obtain all textual data from these PDF documents using a text extraction engine.[3] We also record the date of the meeting and the date and earliest time of the release of each minutes by examining the timestamp of the released file. Our sample is comprised of 196 meeting minutes (hereafter referred to as minutes).

For each minutes, we develop a textual parsing algorithm to do the following: 1) remove the introductory section of the Minutes that lists participant names and administrative matters, and remove the section on specific open market operations (e.g. amount of securities purchased); 2) break the docu-

[2]The delay ranges between three and eight weeks. The Fed implemented a series of accelerated release schedules during the 1990s and 2000s, which shortened the lag from eight (before 2004) to three weeks (after 2004). From 1997 onward, the minutes are released at 2:00pm Eastern Standard Time.

[3]Minutes downloaded in PDF at http://www.federalreserve.gov/monetarypolicy/fomccalendars.htm

ment into individual paragraphs; 3) record the specific section where each paragraph is located (e.g. Staff Economic Discussion or Members' Discussion), and, 4) obtain paragraph length in the number of words. This procedure produces 28,676 unique sentences and 5,644 paragraphs. The average sentence length is 29 words.

C. Market Reaction Data

In many of our tests, we use high-frequency trading data from both equity and bond markets to measure market reactions to the contents of the minutes. We use prices of the exchange-traded fund sponsored by State Street that tracks the S&P 500 index, which we refer to by its ticker symbol SPY, as the proxy for the aggregate stock market. SPY was launched in 1993 and its trading volume has increased dramatically in 2000, making it one of the most liquid instruments. Since volume prior to 2000 is low, we restrict our sample period from 2000 to 2015. As an additional robustness check, we also use proprietary data on the S&P E-Mini futures contracts from the Chicago Mercantile Exchange (CME), which offers similar liquidity levels post-2000. Our results are similar using both instruments.

For the bond markets, we use high-frequency electronic trading data for the Eurodollar futures contracts obtained from the CME. To construct the trading history, we use the "front month" contract, which is the one with expiration dates closest to the date of release of FOMC minutes. Electronic trading was sporadic prior to 2003, ands as a result, we can only construct a reliable trading history of the Eurodollar futures for a shorter sample period from 2003 to 2014.

Next, we construct our event window around the time when the meeting minutes are released. We then calculate return volatility during the event window. The official release time for the meeting minutes is 2:00pm Eastern Standard Time after 1997. Since it is possible that some minutes are released early or late, we search the FOMC's official website, Bloomberg, Dow Jones Newswires, and Thomson Reuters, and we compare the time on the FOMC timestamp with that of the first news story on the day of the Minute's release. We record the release time as the earliest time that the minutes (or news about the minutes) is reported among these sources. The actual release time ranges between 1:59pm and 2:06pm. Therefore, we construct our event window as the 15-minute window between 2:00pm and 2:15pm each day, such that it fully encompasses the reported release times. Our results are robust to alternative event window specifications ranging from 20 minutes to two hours.[4]

[4]We have used windows starting as early as 1:50pm to as late as 2:05pm. We also used window lengths from 20 minutes to

We then calculate event-window return and, following convention, we compute return volatility for the equity market as the squared event-window return. For the bond market, we first compute the LIBOR implied by the Eurodollar future prices (i.e. LIBOR=100-futures price), and yield volatility is the squared value of LIBOR changes. Specifically, for each minutes t in our sample:

$$R_t^{SPY} = \frac{P_{t,2:15pm}^{SPY} - P_{t,2:00pm}^{SPY}}{P_{t,2:00pm}^{SPY}} \tag{1a}$$

$$R_t^{LIBOR} = \frac{Y_{t,2:15pm}^{LIBOR} - Y_{t,2:00pm}^{LIBOR}}{Y_{t,2:00pm}^{LIBOR}} \tag{1b}$$

$$V_t^{SPY} = \left(R_t^{SPY}\right)^2 = \left(\frac{P_{t,2:15pm}^{SPY} - P_{t,2:00pm}^{SPY}}{P_{t,2:00pm}^{SPY}}\right)^2 \tag{1c}$$

$$V_t^{LIBOR} = \left(R_t^{LIBOR}\right)^2 = \left(\frac{Y_{t,2:15pm}^{LIBOR} - Y_{t,2:00pm}^{LIBOR}}{Y_{t,2:00pm}^{LIBOR}}\right)^2. \tag{1d}$$

Since we use a very short window to construct the market volatility measure, confounding effects from other macroeconomic variables are negligible, as the minutes are released predominantly on Wednesdays and (before 2004) Thursdays, and no other significant economic indicators are released on these afternoons.[5] To allow for time-variation in expected volatility, we compute expected volatility as realized volatility of the past 20 days.[6]

II. Methodology

Because the minutes summarize all discussions during the preceding meeting, it contains a wide range of topics. Appendix A presents excerpts from one of the minutes for illustration. As the excerpts indicate, one paragraph discusses the latest developments on inflation, and another paragraph provides the outlook on financial markets. Yet another paragraph discusses both. We examine whether the informativeness of the FOMC minutes vary across these topics. This section describes our methodology to separate the FOMC minutes into individual topics and extract the content from each topic.

2 hours, in 10-minute increments. The results are similar throughout most window lengths.

[5] See http://www.bloomberg.com/markets/economic-calendar for a schedule of important economic news. Usually no other significant news is scheduled to release on Wednesdays. On Thursdays, most other indicators are released on the morning prior to market open.

[6] We varied the volatility estimation window from past 1 to 5 days to past 1 to 30 days. The results are not sensitive to these changes in the estimation window.

A. *The Latent Dirichlet Allocation (LDA) Algorithm*

We use the Latent Dirichlet Allocation (LDA) algorithm first developed by Blei, Ng, and Jordan (2003) to classify each minutes into distinct topics. The LDA belongs to a broader class of probabilistic topic models that use hierarchical Bayesian analysis to uncover the underlying semantic structure of textual documents. These models use two statistical distributions to describe the latent data generating process for each paragraph, which is the base unit of our analysis.

The LDA summarizes each paragraph as a distribution over a collection of topics, each of which is in turn a distribution over the collection of English words that the sample texts use. For example, a paragraph that discusses inflation would be represented by a distribution that places large weights on words such as prices, CPI, inflation, etc. By contrast, a topic that discusses foreign trade would place large weights on words such as trade and imports, but a small weight on words such as CPI and inflation.

Since these two distributions are latent, these methods use Bayesian techniques to efficiently estimate the parameters of the unobservable distributions from the data (i.e. the collection of documents). LDA represents one particular parameterization of the model that assumes that the latent distributions belong to the Dirichlet family. Then, armed with this functional form and the observed words in each paragraph, we compute the posterior distributions for each paragraph and topic using the standard Bayes Theorem. These empirical posterior distributions are the main outputs of the model. The only inputs in LDA are the document texts and the number of topics. Since it requires no training data or a priori tuning of parameters, the LDA belongs to the class of unsupervised machine learning models.

We illustrate our approach with a simple example. Suppose we are given the following three paragraphs:

1. *The employment situation is good and layoffs have declined.*
2. *Imports have increased and the outlook for trade is good.*
3. *Imports look good, and the employment situation is also good.*

For the purposes of illustration, suppose the full set of relevant FOMC vocabulary consists of only $V = 4$ words {*employment*, *layoffs*, *imports*, *trade*}.

A human reader would intuitively recognize that the first paragraph is about employment and the second is about foreign trade. The third paragraph is a mixture of both. Suppose we fit the LDA model with $N = 2$ topics. If the model performs satisfactorily, then the posterior topic distributions should

clearly determine the weights for each word that would define each topic. For instance, the output would be something similar to:

- $\hat{\beta}_1 \equiv \{\hat{P}_{topic1}(employment), \hat{P}_{topic1}(layoff), \hat{P}_{topic1}(imports), \hat{P}_{topic1}(trade)\}$
  $= \{0.55, 0.43, 0.01, 0.01\}$
- $\hat{\beta}_2 \equiv \{\hat{P}_{topic2}(employment), \hat{P}_{topic2}(layoff), \hat{P}_{topic2}(imports), \hat{P}_{topic2}(trade)\}$
  $= \{0.01, 0.01, 0.60, 0.48\}$,

where $\hat{\beta}_i$ is the vector of probabilities $\hat{P}_j$ of observing word j in topic i. Next, vectors $\hat{\theta}_i$ characterize posterior distribution of the probabilities that paragraph i is about topic j. In our example, if the LDA output is consistent with human intuition, these vectors would have values similar to the following:

- $\hat{\theta}_1 \equiv \{\hat{P}_{paragraph1}(Topic1), \hat{P}_{paragraph1}(Topic2)\} = \{0.99, 0.01\}$
- $\hat{\theta}_2 \equiv \{\hat{P}_{paragraph2}(Topic1), \hat{P}_{paragraph2}(Topic2)\} = \{0.01, 0.99\}$
- $\hat{\theta}_3 \equiv \{\hat{P}_{paragraph3}(Topic1), \hat{P}_{paragraph3}(Topic2)\} = \{0.51, 0.49\}$

B. Choosing the Number of Topics and LDA Implementation

The LDA classification of the FOMC minutes generalizes this example to our sample of $D = 5,644$ unique paragraphs. This set of paragraphs constitutes our corpora and the input to the LDA algorithm. Stop words, such as a, the, etc., are removed prior to processing.[7] This results in a collection of $V = 61,432$ words.

Next, we need to choose the number of topics, which is typically chosen by the researcher based on the context. To guide our choice of number of topics, we fit a modified version of the LDA, where the number of topics is also a parameter to be estimated from the data.[8] For computational tractability, we fit the algorithm 250 times, each with 300 randomly selected paragraphs from our corpora. We report the summary statistics from this exercise in the Online Appendix. The results show that the number of topics estimated from the data is centered around 8. Based on this analysis, we choose $N = 8$ in our empirical analysis.[9]

[7] The stop words we use are available as part of the Online Appendix at https://michiganross.umich.edu/faculty-research/faculty/andrew-wu.

[8] The model is based on the Hierarchical Dirichlet Process model from Blei et al. (2004).

[9] Our results are robust to alternative specifications from $N = 5$ to $N = 10$. When we allow for a larger number of topics, some topics become redundant. However, the algorithm results in a similar number of major topics after we group similar topics together.

Each of the N topics represents a distribution over the V words in the FOMC vocabulary, and each paragraph is a mixture of the N topics. We assume that the observable data, i.e. words in each document, is generated from a probabilistic data generating process parameterized as follows:

1. Each of paragraph $d = 1, \ldots, D$ contains a mixture of N topics. Let the proportion of topic n in paragraph d be $\theta_{d,n}$ and let the vector $\theta_d = [\theta_{d,1}, \ldots, \theta_{d,N}]'$ represent the true topic mixture of paragraph d. For each d, we assume that this mixture follows an order-N Dirichlet distribution over the N topics, governed by the latent parameter vector μ of size N:

$$\theta_d \sim Dirichlet_N(\mu).$$

2. Given paragraph d's topic mixture θ_d, let the assignment of each word i in document d into topics be $Z_{d,i}$, where $Z_{d,i} \in \{1, \ldots, N\}$. We assume that this assignment follows the multinomial distribution governed by the document-specific topic vector θ_d described in the previous step:

$$Z_{d,i}|\theta_d \sim Multinomial(\theta_d). \tag{2}$$

   Suppose there are I_d unique words in document d. Let the vector Z_d denote the collection of the topic assignment of all words within d, i.e. $Z_d = \{Z_{d,i}\}_{i=1}^{I_d}$

3. The N topic distributions (applied universally to all paragraphs) are in the collection $\beta = \{\beta_1, \ldots, \beta_N\}$. Each topic β_n also follows an order-V Dirichlet distribution over the V words, governed by the latent scalar parameter ϕ:

$$\beta_n \sim Dirichlet_V(\phi). \tag{3}$$

4. For each word i in document d, there are V choices to choose from based on our FOMC vocabulary. Conditional on the chosen topic for word i in Step 2 above (i.e. a draw from Distribution 2), and on the structure of the topic distribution from Step 3 (i.e. a draw from Distribution 3), we assume that actual choice of the word, $W_{d,i}$, follows a multinomial distribution governed by the resulting word-topic assignment $\beta_{Z_{d,i}}$:

$$W_{d,i}|\left(\{\beta_n\}_{n=1}^{N}, Z_{d,i}\right) \sim Multinomial(\beta_{Z_{d,i}}).$$

Similarly, let the W_d denote the collection of the vocabulary choice of all words within document d: $W_d = \{W_{d,i}\}_{i=1}^{I_d}$.

The above four distributions constitute the latent data generating process that results in our observable document collection $\{W_d\}_{d=1}^{D}$. Recall that they are not directly observable to the researcher. Instead, the only observable data is the occurrence of the actual words i in each document d, i.e. W_d. We can then write the overall data generating process as the joint distribution of latent variables $\{\beta_n\}_{n=1}^{N}, \{\theta_d\}_{d=1}^{D}, \{Z_d\}_{d=1}^{D}$ and the observable variable $\{W_d\}_{d=1}^{D}$:

$$\begin{aligned} &P\left(\{\beta_n\}_{n=1}^{N}, \{\theta_d\}_{d=1}^{D}, \{Z_d\}_{d=1}^{D}, \{W_d\}_{d=1}^{D}\right) \\ &= \prod_{n=1}^{N} P(\beta_n) \prod_{d=1}^{D} P(\theta_d) \left[\prod_{i=1}^{I_d} P(Z_{d,i}|\theta_d) P\left(W_{d,i}|\{\beta_n\}_{n=1}^{N}, Z_{d,i}\right) \right], \end{aligned}$$

where $P(\cdot)$ are the respective (Dirichlet or multinomial) density functions specified above.

Now that we observe our FOMC document collection $\{W_d\}_{d=1}^{D}$, we can compute the posterior distribution of the document-topic structure given the observed documents using Bayes' Rule:

$$P\left(\{\beta_n\}_{n=1}^{N}, \{\theta_d\}_{d=1}^{D}, \{Z_d\}_{d=1}^{D} | \{W_d\}_{d=1}^{D}\right) = \frac{P\left(\{\beta_n\}_{n=1}^{N}, \{\theta_d\}_{d=1}^{D}, \{Z_d\}_{d=1}^{D}, \{W_d\}_{d=1}^{D}\right)}{P\left(\{W_d\}_{d=1}^{D}\right)}. \tag{4}$$

Similar to other Bayesian inference methods, the numerator in Equation (4) can be easily computed, but the denominator is by construction a double integral and therefore cannot be feasibly computed. However, it can be efficiently approximated using a standard Gibbs sampler.

C. Results from the LDA Inference

Once the posterior probabilities are computed, we compute the posterior expectations of two key latent variables, which represent the main output from the LDA algorithm:

1. Posterior vocabulary distribution for each topic: $\{\hat{\beta}_1, \ldots, \hat{\beta}_N\}$
2. Posterior topic mixture for each paragraph in our collection: $\{\hat{\theta}_1, \ldots, \hat{\theta}_D\}$

The first set of output from our LDA procedure identifies frequency of occurrence of each word in each of the topics. For each topic k, $\hat{\beta}_k = [\hat{\beta}_{k,1}, \ldots, \hat{\beta}_{k,V}]'$ and each entry $\hat{\beta}_{k,j}$ represents the probability that the word j characterizes topic k. Our FOMC document collection has 61,432 unique words.

Therefore, each $\hat{\beta}_k$ contains 61,432 entries, the majority of which receives a weight close to zero. Table I reports the top 20 words for each topic. We use the top words to identify the topic.

The top words from each classified topic are mostly distinct and identify their respective topics with little ambiguity. For example, the first topic consists of keywords such as *policy*, *stance*, *easing* and *tightening*, indicating that this topic is about monetary policy related to interest rate targeting and management. The second topic consists of keywords such as *inflation*, *energy*, etc., indicating that this topic is about inflation. In fact, the rest of the topics can be similarly identified by the top keyword from their respective classification, as 3) financial market, 4) employment, 5) economic growth, 6) foreign trade, 7) consumption, and 8) production and investment.

The second set of output is the collection of paragraph-level topic mixture vectors, $\{\hat{\theta}_1, \ldots, \hat{\theta}_D\}$. From this collection, each paragraph d has one mixture, $\hat{\theta}_d = [\hat{\theta}_{d,1}, \ldots, \hat{\theta}_{d,N}]'$. Because there are eight topics, each vector $\hat{\theta}_d$ has eight entries, where each $\hat{\theta}_{d,n}$ corresponds to the proportion of paragraph d that is devoted to topic n. The eight entries sum up to one for each paragraph. We plot the time series of the proportion of each topic in Figure 1. The shaded area in Figure 1 corresponds to NBER-designated recession periods.

Figure 1 shows significant time variation in the proportion of the FOMC minutes devoted to each topic. For example, a progressively smaller portion of the minutes is devoted to the growth topic over time, which was the most discussed topic in the early part of the sample period.

The topics that gain a larger share during this period are policy stance and market. This pattern suggests that the Fed is increasingly focused on its regulatory role in maintaining the stability of the financial markets, such as negotiating the rescue of systemically important banks and the subsequent TARP initiatives. The proportion of the minutes devoted to financial markets tripled during the recent financial crisis.

We next examine whether the proportion of various topics that FOMC discusses varies with the macroeconomic conditions using the following proxy variables for the state of the economy:

- $IntRate$: the latest daily closing yield of 10-year Treasury notes obtained from the Federal Reserve Board's H.15 releases.
- $UnEmp$: latest monthly rate of unemployment obtained from the Bureau of Labor Statistics.
- $Recession$: a dummy variable which equal to one if meeting date t falls within a NBER-designated

recession period.

We fit the following regression to examine the relation between the proportion of minutes devoted to various topics and macroeconomic conditions:

$$\hat{\theta}_{n,t} = a + b \times IntRate_t + r \times UnEmp_t + d \times Recession_t + e_t, \tag{5}$$

where $\hat{\theta}_{n,t}$ is the topic-n proportion in Minutes t estimated from the LDA procedure. We fit the regression using all 196 minutes from 1991 to 2015.

[Insert Table II about here]

Table II presents the coefficient estimates. The level of interest rates is positively related to discussion about growth, investment and trade, and negatively related to inflation, market and consumption. Unemployment is positively related to market and negatively related to several other topics including employment. It appears that the Fed focuses more on how the markets influence employment than on direct employment related topics per se when unemployment is high. The Fed also devotes a larger proportion of its discussions on markets during recessions than on other topics. Overall, these results indicate that proportion of various topics vary with macroeconomic conditions.

D. Extraction of Contents

We measure the tone of each paragraph using a bag-of-words approach similar to Tetlock (2007), Loughran and McDonald (2011) and Jegadeesh and Wu (2013). Specifically, for each paragraph, we compute the tone of each topic by tabulating the frequency of keywords in the respective tone lexicons. We compute the paragraph tone as the difference between the number of positive and negative tonal words. We use the comprehensive tonal lexicon that merges the Harvard IV-4 Psychosociological Dictionary[10] and the financial tonal lists developed by Loughran and McDonald (2011) to categorize the tone of each word.[11] A higher Tone Score indicates a more positive/easing stance or less negative/tightening tone.

Next, we aggregate the Scores to the document level as the sum of individual paragraph scores, weighted by the inverse of paragraph length in number of words:

[10] Available at http://www.wjh.harvard.edu/∼inquirer/homecat.htm.
[11] Available at http://www3.nd.edu/∼mcdonald/Word_Lists.html.

$$Score_{n,t} = \sum_{d=1}^{D_t} Score_{d,n,t} \left(\frac{1}{T_d^t} \right), \tag{6}$$

where T_d^t is the total number of words in paragraph d, and D_t is the total number of paragraphs in Document t. The term $\left(\frac{1}{T_d^t} \right)$ reflects the intuitive notion that the strength of the topic tone is negatively related to overall paragraph length. Longer paragraphs are more difficult to read and process, and they are therefore down-weighted.

Figure 2 plots the 10-period moving average of the document-level Tone Scores of each of the eight topics over time, from the March 1992 meeting to the June 2015 meeting. For ease of comparison, we use standardized Scores by subtracting their respective time-series mean and dividing by their respective standard deviations. The top and bottom panels of Figure 2 display the Tone Scores for topics 1-4 and 5-8, respectively. The tone of the topics shows systematic variation over time. The tone scores are generally procyclical, becoming more positive during boom periods and turning sharply negative during recessions. The figure presents suggestive evidence that the tone of the policy topic leads economic upturn during recessions, as it seems to turn halfway into the recession, even before the tone of other topics changes direction. However, it is hard to draw any strong conclusion because there are only two recessions during our sample period.

We examine whether the tone of the topics depend on economic conditions using the following regression:

$$Score_{n,t} = \alpha + \beta IntRate_t + \gamma UnEmp_t + \delta Recession_t + \epsilon_t, \tag{7}$$

where $Score_{n,t}$ is the Net Tone Content Score for topic n and Minutes t, computed per Equation (6). Table III reports the coefficient estimates for the Tone Score regressions.

The most important determinant of the tone is the recession dummy. During recessions, the tones of discussion of all topics tend to be more negative. The relations between unemployment and tone of a number of topics, including market, employment and economy, are negative. Both these results indicate that the tone of the FOMC minutes reflect negative economic sentiments during troubled times.

III. Empirical Tests and Results

This section examines the informativeness of FOMC minutes. We first assess the informativeness of the minutes as a whole, and then we measure the relative informativeness of each individual topic using contemporaneous market reaction.

A. *Data*

We examine the relation between the content of FOMC minutes and changes in aggregate stock market returns and interest rates to assess the information content of FOMC minutes. We use transaction prices of SPY to measure intraday market returns. We obtain SPY transaction price data from TAQ. We use the 3-month LIBOR rate implied by the nearest maturity Eurodollar futures contract as the interest rate measure.[12] We obtain transaction prices of Eurodollar futures from the Chicago Mercantile Exchange.

B. *Informativeness of FOMC Minutes as a Whole*

To assess the informativeness of the minutes, we first examine the market reaction soon after the Fed releases the minutes, using the following regression specifications:

$$V_t = \alpha + \beta L_t + \sum_{k=1}^{K} \gamma_k V_{t-k} + \epsilon_t, \tag{8}$$

where V_t is the 15-minute event window market volatility (on both release and non-release days) computed per Equations (1c) and (1d). L_t is a dummy variable that equals one if a minutes document is released on date t. We use V_{tk}, the k-day lagged daily volatilities as control variables to account for time-varying volatilities. We fit Regression (8) above using K=0, 5, 10, and 20 trading days. The Minutes are released at 2:00 pm and hence we use the 15-minute window from 2:00 pm to 2:15 pm as the event window. To facilitate interpretation, we scale all regression coefficients by the unconditional mean of V_t across all observations. The coefficient estimate $\hat{\beta}$ can thus be interpreted as the incremental volatility introduced by the release of the minutes as a fraction of the average volatility in the event window across both release and non-release days. Each regression uses between 4,343 and 4,363 observations.

[12] Implied 3-month LIBOR=100-Eurodollar futures price.

[Insert Table IV about here]

Table IV reports the coefficient estimates. The estimate for the release dummy, L_t, is significantly positive for all specifications, and it ranges from 0.5919 to .6130 for SPY. The inclusion of lagged volatility as control variables increases regression R^2 since it accounts for time-variation in volatility, but it does not materially affect the slope coefficients. These estimates indicate that when the minutes are released the volatility is about 60% higher than that during the same time on other days.

Table IV also reports the results for volatility of LIBOR. The slope coefficients for LIBOR are between .2283 and .3054. The proportional increase in interest rate LIBOR volatility is about half that for SPY, indicating that the minutes have a relatively larger impact on the stock market.

[Insert Figure 3 about here]

Figure 3 also examines the volatility changes following the release of the minutes. We plot the ratio of standard deviation of percentage changes in SPY prices and LIBOR within 15-minute intervals on release days relative to that during the corresponding time on non-release days. The ratio fluctuates around one until 2 pm. It then spikes to about 1.6 for SPY and 1.3 for LIBOR at 2 pm. The ratio declines to 1 by 2:30 pm, which indicates that the information in the minutes is quickly incorporated into asset prices.

Interestingly, the ratio dips below one at 2:30pm and stays below one the rest of the day. The decline in volatility after 2:30 pm indicates that the market receives less information during this period on release days than on other days. Together, these findings indicate that some of the information that reaches the market through the minutes during the 2:00 PM to 2:30 PM window would have reached the market anyway before the market close on release days. However, the volatility during the 2:00PM to market close window is larger on release days than on non-release days (averaging 1.13 than on the other days for SPY and 1.07 times for LIBOR), indicating that the shift in timing of information is not the entire reason for the spike in volatility immediately after the release of the minutes.

C. Information in Tone of FOMC Minutes: Full Document

The increased volatility during the release of the minutes indicates that the FOMC conveys information to the market. This section examines whether the composition of topics discussed in the meetings and the tone of the discussion convey incremental information.

Our first set of tests considers the tone of the entire document and our subsequent tests consider the tones of individual topics. We use the following regression in our tests:

$$UV_t = \alpha + \beta_c Score_c^t + \epsilon_t \tag{9}$$

where UV_t is the unexpected volatility around the event window on release date t, as the raw volatility minus the average market volatility in the past 20 days. That is, given V_t computed from Equation (1c) or (1d):

$$UV_t = V_t - \frac{\sum_{k=1}^{21} V_{t-k}}{20}. \tag{10}$$

Our main tests use the Tone Scores as explanatory variables. We present the results using separate Positive and Negative Tone Scores in an Online Appendix, which are qualitatively similar. The results in the last section indicate that content of the FOMC minutes are related to various macroeconomic variables. Therefore, we use them as control variables in the regressions.

[Insert Table V about here]

Table V reports the regression estimates. The slope coefficient for SPY and LIBOR are -0.0159 and 0.0037. The coefficient on tone is marginally significant for SPY, and it is not statistically significant for LIBOR. These results indicate that the stock market finds the tone of the entire minutes marginally more informative when it is negative than when it is positive, but the tone is not incrementally informative for the debt market.

Among the control variables, the slope coefficient on the recession dummy for the SPY regression is significant, but the other slope coefficients are not reliably different from zero. None of the slope coefficients on the control variables is significant in the LIBOR regression. Overall, these results indicate that the stock market finds the minutes marginally more informative during recessions, but the informativeness of the minutes do not seem to significantly vary with economic conditions.

D. Informativeness of Individual FOMC Topics

This section examines whether the proportion of the minutes that the Fed devotes to individual topics and the tone of these topics are incrementally informative to the market using the following

regressions:

$$UV_t = \sum_{n=1}^{8} b_n \times \hat{\theta}_{n,t} + r \times X_t + e_t \tag{11a}$$

$$UV_t = \alpha + \sum_{n=1}^{8} \beta_n \times Score_{c,n}^{t} + \gamma \times X_t + \epsilon_t \tag{11b}$$

where UV_t is the unexpected volatility around the event window on release date t, computed per Equation (10). We fit the regressions for both SPY and LIBOR, and we include the control variables for economic conditions.

The sample period for SPY is from 2000 to 2015 and for LIBOR is from 2003 to 2014. In this setting, an estimate of b_n or $\beta_{c,n}$ that is statistically different from zero would indicate incremental informativeness of a topic, or its content Score. Specifically, a significantly positive β would suggest that the market responds more to a more positive topic tone while a significantly negative β would suggest that the market finds more information in a more negative topic tone. Similarly, a significantly positive b for topic n would indicate that the market finds the discussion of this topic informative when it is discussed more, regardless of the tone.

[Insert Table VI about here]

Table VI presents the coefficient estimates from the proportion Regression (11a). Because the topic proportions sum up to one in all documents, we fit this regression without an intercept to prevent multicollinearity. All independent variables in the regressions are standardized to mean zero and unit standard deviation.

In the unexpected volatility regressions, the slope coefficients on policy, inflation and employment are significantly positive for both SPY and LIBOR. For SPY, investment is also significantly positive. The other topics are not incrementally informative. These findings make intuitive sense because the actions of the Fed would likely have more direct impact on issues related to these topics.

We also examine whether the information in proportions of topics discussed moves the market in one direction or another. We use the raw returns R_t computed from Equation (1a) or changes in interest rates as the dependent variables in place of unexpected volatility in Regressions (11a) and (11b). Table VI reports the results.

The slope coefficient on policy stance is statistically significant for the SPY with controls variables for economic conditions and for LIBOR with and without controls. The coefficient is positive for equity

and negative for interest rates. Therefore, when the Fed devotes more attention to its policy stance in the minutes, it seems to take an easing stance that results in a reduction in interest rates, which is also viewed as good news for the stock market. The coefficient on employment is significantly negative for interest rates. A larger proportion of the minutes devoted to employment probably reflects Fed's concerns about this topic. Therefore, the market seems to expect lower rates in the future. Here again, the market seems to be influenced more by the Fed's discussion of topics over which it has direct control.

[Insert Table VII about here]

Table VII presents the results of tone regressions (11b). For SPY, the slope coefficients on the tone of inflation and employment are significantly negative, indicating that the minutes are more informative when the tone is negative than when it is positive. The slope coefficients on growth and consumption are positive. The other slope coefficients are not reliably different from zero. For LIBOR, the slope coefficients on policy, inflation and employment are significantly negative. These results indicate that the debt market finds a negative tone on these topics more informative than a positive tone.

We also examine the directional impact of tone of each topic. The slope coefficient on the tone of the policy is significantly positive for SPY and significantly negative for LIBOR. Therefore, when the Fed takes up an easing posture, the stock market tends to go up and the interest rate tends to go down. We find a similar result for inflation as well. None of the other topics are significantly related to the directional move of LIBOR. However, the tone of the market, economy and investment are negatively related to the direction of SPY. These results indicate that the market interprets the Fed's positive tone on these topics to mean that the Fed is unlikely to take actions to help boost the market.

Overall, these results indicate that some topics of the FOMC minutes are more informative to the market than others. Also, in the case of SPY, the slope coefficients for some topics are significantly positive while some others are significantly negative. These finer details about the incremental information content are not evident when we examine the document in its entirety.

IV. Alternative Specifications and Robustness Checks

A. *Subsample Evidence*

FOMC minutes are longer after 2011 than before.[13] For example, minutes released before 2011 contain on average 3,863 words and 41.19 paragraphs, while those released after 2011 contain an average of 5545 words and 25.91 paragraphs. This section examines whether the informativeness of the minutes are different in the recent subperiod from that in the shorter subperiod.

[Insert Figure 4 and Table VIII about here]

Figure 4 presents the relative standard deviations of returns during the 2:00pm-2:15pm window from days t-3 to t+3 during the full sample of 2000-2015, as well as with the 2000-2010 and 2011-2015 subsamples. We normalize the volatility at t-3 to 1, and plot the three series. Consistent with our results on unexpected volatilities, realized volatilities on the announcement dates are higher than that on non-announcement dates in the full sample and both subsamples. The stock market volatility on the release date is about 250% of that during the pre-release window in the post-2011 period and about 150% in the pre-2011 period. For LIBOR, the volatility on the release date is about 130% of that during the pre-release window in both subperiods. The results indicate that the more detailed minutes in the recent subperiod are more informative for the stock market but not for the bond market.

B. *Results for Meetings With and Without Summary of Economic Projections*

Beginning with the October 2007 meeting, all participants (including 7 members of the Federal Reserve Board and 12 Presidents of regional Federal Reserve Banks) submit individual economic projections in conjunction with four out of eight meetings each year. The summary of these economic projections (SEP) is included as an addendum to the minutes. Moreover, from April 2011 onwards, an advance version of the SEP data is released at the conclusion of the meetings in conjunction with the Chairman's press conference.[14] The key component of the SEP data is the median, central tendency, and the range of forecasts on real GDP growth, unemployment rate, inflation, and the Federal funds rate, up to five years in horizon. Since the SEP contains numerical data in addition to the text, it is

[13] Fed personnel involved with preparing the minutes noted during our discussions that the Fed minutes since 2011 are longer and more detailed, which we confirmed in the data.

[14] See https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm#sep for more information.

possible that the minutes with the SEP may be more informative than minutes without the SEP. Since the SEP is only released in four out of eight meetings per year, these releases represent an exogenous environment where some minutes' topics might be less informative.

[Insert Figure 5 and Table IX about here]

Figure 5 presents the relative standard deviations of returns during the 2:00pm-2:15pm window from days t-3 to t+3 for the 31 minutes with concurrent SEP releases from 2007 to 2015 and the 33 minutes without the SEP. As before, we normalize the volatility on day t-3 to 1. These results indicate that the volatility spikes on the release dates are about the same for both minutes with the SEP and without the SEP. Therefore, much of the informativeness of the minutes is attributable to its textual contents.

We also fit Regressions (11a) and (11b) within these subsamples separately and Table IX reports the estimates. The topics that are significant for the full sample in Table VII by and large continue to be significant in each of the subsamples. We also did not find any noticeable difference in the informativeness of the individual topics between these subsamples. Overall, we do not find any evidence that the announcement of the SEPs has a noticeable change in informativeness of the minutes.

C. Alternative Lexicons and Tone Measures

Our tests so far use a comprehensive lexicon that merges the LM word list with Harvard IV-4 Psychosociological Dictionary. Since both these lexicons are subjectively created, we examine the robustness of our results to the choice of lexicons. Specifically, we compute tone scores using the LM word list and the Harvard IV-4 Psychosociological Dictionary separately and fit Regressions (11a) and (11b) with these scores, with the full set of controls.

Table X reports the results for the tone regressions for SPY and LIBOR. For the SPY regression, the slope coefficients are about the same with both lexicons. For example, inflation, employment and growth are statistically significant. The slope coefficient on consumption is marginally significant with IV4 lexicon, but it is not significant with the LM lexicon. However, these coefficients are not statistically different. All of the other slope coefficients are not significant. None of the slope coefficients with individual lexicons are statistically different from the corresponding coefficients in Table VII that uses the combined lexicon. The LIBOR regressions also yield similar results.

In the directional price change regression, all slope coefficients, except those on growth and trade, are statistically significant with the LM lexicon, which is similar to the results we find in Table VII using the combined lexicon. The signs of all the slope coefficients with the LM lexicon are also the same as the corresponding coefficients with the combined lexicon and their magnitudes are similar. We also find similar results with the IV4 lexicon, except for the consumption slope coefficient. This coefficient is not statistically significant but it is also not different from the corresponding coefficients with the other lexicons. These results indicate that our approach is robust with respect to the choice of alternative lexicons.

[Insert Table X about here]

Our approach so far follows common practice and assigns equal weights to all tonal words to compute the tonal score. We now consider other approaches that weight the tonal words based on their likely impact. A commonly used approach is to weight each word inversely proportionally to their frequency of occurrence in the document sample. This weighting scheme captures the idea that one may expect that words that occur less frequently are more likely to be context specific and hence deserve more weight. These weights are called the tf.idf weights (see Manning and Schütze, 1999.) Jegadeesh and Wu (2013) propose an alternate weighting scheme that assigns weights based on the impact of tonal words used in 10-Ks. They compute what they call word-power weights based on the relation between market reactions to the filing of 10-Ks and the tonal words used in those documents.

Table X presents the estimates of Regressions (11a) and (11b) with tone scores computed with tf.idf weights as well as with word power weights. For both the volatility and price change regressions, the policy, market, growth, and investment topics remain statistically significant, similar to the results in Table VII. The inflation topic remains significantly negative for the volatility regression, but is not statistically significant for the price change regression. Overall, these results indicate that our approach is robust to different tonal word weighting schemes.

V. Conclusion

We present a new approach to quantify the economic and policy content of Federal Reserve communications by dissecting the Federal Open Market Committee (FOMC) meeting into distinct economic

topics. Our approach uses the Latent Dirichlet Allocation (LDA) algorithm to determine the topics discussed in each paragraph of the minutes. We examine the informativeness of the proportion of the minutes devoted to each topic and the tone of the topics for the stock and bond markets. We use SPY, an actively traded ETF that tracks the S&P 500 index, and the 3-month LIBOR implied by Eurodollar futures contract prices as proxies for the stock and bond markets, respectively.

We identify the following eight topics in the minutes: policy stance, inflation, financial market, employment, economic growth, foreign trade, consumption and production and investment. We find significant variation in the proportion of various topics that the FOMC discusses and the tone of the topics over time. The tone of the discussion tends to be negative during recessions and when unemployment is high. We also find that the contents of some topics are more informative than others. For example, discussions of policy stance and inflation are informative, but discussions of other topics such as trade and consumption are not.

We also examine the directional impact of the tone of each topic. We find that the tone of policy stance is significantly positive for SPY, but significantly negative for LIBOR. These results indicate that when textual content of the minutes imply that the Fed will take an easing posture, the stock market tends to go up and the interest rate tends to go down. We find a similar result for inflation. Our results are robust within subperiods and also with respect to changes in the choice of tonal lexicons and in the number of topics.

Our analysis focuses on the relation between the textual content of the FOMC minutes and contemporaneous reactions of the prices of financial assets in the stock and bond markets. Contemporaneous price changes reflect revisions in market expectations about the real economy based on the information conveyed by the text of the FOMC minutes. An examination of the direct relation between the textual content of the minutes and the real economy is likely a fruitful avenue for future research.

REFERENCES

Ang, Andrew, Jean Boivin, Sen Dong, and Rudy Loo-Kung, 2002, Monetary policy shifts and the term structure, *Review of Economic Studies* 78, 429–457.

Blei, David M, Andrew Y Ng, and Michael I Jordan, 2003, Latent dirichlet allocation, *Journal of Machine Learning Research* 3, 993–1022.

Campbell, John Y., Carolin Pflueger, and Luis M. Viceira, 2015, *Monetary policy drivers of bond and equity risks* (Working paper, Harvard University, Cambridge, MA).

Fleming, Michael J., and Monika Piazzesi, 2005, *Monetary policy tick-by-tick* (Working paper, University of Chicago, Chicago, IL).

Friedman, Milton, and Anna Jacobson Schwartz, 1963, *A monetary history of the United States, 1867-1960* (Princeton University Press).

Hanley, Kathleen Weiss, and Gerard Hoberg, 2010, The information content of IPO prospectuses, *Review of Financial Studies* 23, 2821–2864.

Jegadeesh, Narasimhan, and Di Wu, 2013, Word power: A new approach for content analysis, *Journal of Financial Economics* 110, 712–729.

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance* 66, 35–65.

Lucca, David O., and Emanuel Moench, 2015, The pre-FOMC announcement drift, *Journal of Finance* 70, 329–371.

Manning, Christopher D, and Hinrich Schütze, 1999, *Foundations of statistical natural language processing*, volume 999 (MIT Press).

Primiceri, Giorgio E, 2005, Time varying structural vector autoregressions and monetary policy, *The Review of Economic Studies* 72, 821–852.

Romer, Christina D, and David H Romer, 1989, Does monetary policy matter? a new test in the spirit of Friedman and Schwartz, in *NBER Macroeconomics Annual 1989, Volume 4*, 121–184 (MIT Press).

Rosa, Carlo, 2013, The financial market effect of FOMC minutes, *FRBNY Economic Policy Review* 19.

Sargent, Thomas, Noah Williams, and Tao Zha, 2006, Shocks and government beliefs: The rise and fall of American inflation, *The American Economic Review* 96, 1193–1224.

Sims, Christopher A., and Tao Zha, 2006, Were there regime switches in U.S. monetary policy?, *American Economic Review* 96, 54–81.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62, 1139–1168.

Appendix A Select FOMC Paragraphs with LDA Classification Results

Part 1. Single Topic Examples

Example 1. (99% growth mandate, other topics negligible)

With regard to developments and prospects in key sectors of the economy, members noted that despite further survey indications of eroding consumer confidence, consumer expenditures had strengthened in recent months after a pause earlier in the year. The pickup had featured rising sales of motor vehicles, and while the latter had slipped recently, a number of special factors such as shortages of popular models at the end of the model year and the effects of flooding in some parts of the Midwest suggested the need to withhold judgment on any downward shift in the underlying demand for motor vehicles. Tourism was reported to have strengthened considerably in many areas this summer, though there were major exceptions. As had been true for an extended period, consumer attitudes continued to be inhibited by concerns about employment opportunities, especially given further reductions in defense spending, the ongoing restructuring and related downsizing of many business operations, and the continuing efforts by business firms to limit the number of their permanent employees in order to hold down the rising costs of health care and other nonwage worker benefits. Members noted, however, that the growth in employment thus far this year, while tending to involve many low paying jobs, had greatly exceeded the rate of expansion in 1992. In the view of at least some members, appreciable further growth was likely as business firms found it increasingly difficult in an expanding economy to meet growing demands through outsourcing, temporary workers, and overtime work. Some members also noted that the newly legislated taxes on higher incomes would tend to curtail some consumer spending. The timing of that effect was uncertain; tax liabilities had already risen, but some payments on the added tax liabilities were not due until April of 1994 and 1995.

Example 2. Inflation Mandate (99% inflation mandate, other topics negligible)

The core consumer price index advanced at a faster rate in the first quarter than it had in the fourth quarter, reflecting the pass-through of higher energy prices and a leveling off of goods prices after sizable declines last year. The higher goods price inflation owed, in part, to the recent run-up in the prices of non-oil imports, energy, and other commodities. The price index for core personal consumption expenditures also rose at a faster rate in the first quarter than it had late last year. Despite the rise in inflation this year, however, the cumulative increase in the overall consumer price index for the year ending in March was somewhat less than the advance for the twelve months ending in March 2003. In the year ending in March, the increase in the price index for total personal consumption expenditures was similar to that of a year earlier. Survey measures of near-term inflation expectations edged up somewhat in March and April, but measures of longer-term expectations decreased. With regard to labor costs, average hourly earnings of production or nonsupervisory workers on private nonfarm payrolls rose notably less for the twelve months ending in March than they had in the year-earlier period. The overall increase in the employment cost index for private industry for the twelve months ending in March was about the same as that for the twelve-month period ending a year earlier, as wages and salaries decelerated and benefits accelerated.

Example 3. Financial Market Mandate (99% market mandate, other topics negligible)

Participants noted that financial markets were volatile over the intermeeting period, as investors responded to news on the European fiscal situation and the negotiations regarding the debt ceiling in the United States. However, the broad declines in stock prices and interest rates over the intermeeting period were seen as mostly reflecting the incoming data pointing to a weaker outlook for growth both in the United States and globally as well as a reduced willingness of investors to bear risk in light of the greater uncertainty about the outlook. While conditions in funding markets had tightened, it was noted that the condition of U.S. banks had strengthened in recent quarters and that the credit quality of both businesses and households had continued to improve.

Example 4. Policy Mandate (99% policy mandate, other topics negligible)

Participants discussed a number of policy tools that the Committee might employ if it decided to provide additional monetary accommodation to support a stronger economic recovery in a context of price stability. One of the policy options discussed was an extension of the period over which the Committee expected to maintain its target range for the federal funds rate at 0 to 1/4 percent. It was noted that such an extension might be particularly effective if done in conjunction with a statement indicating that a highly accommodative stance of

monetary policy was likely to be maintained even as the recovery progressed. Given the uncertainty attending the economic outlook, a few participants questioned whether the conditionality of the forward guidance was sufficiently clear, and they suggested that the Committee should consider replacing the calendar date with guidance that was linked more directly to the economic factors that the Committee would consider in deciding to raise its target for the federal funds rate, or omit the forward guidance language entirely.

Part 2. Multiple Topic Examples

Example 5. (56% growth, 43% inflation)

The information reviewed at this meeting suggested that economic activity had weakened further in the opening months of the year. Production cutbacks were evident in a wide range of industries, and private payrolls had fallen markedly, especially in the goods producing sector. On the positive side, consumer confidence had rebounded sharply since the cease-fire in the Persian Gulf, retail sales and housing starts had strengthened recently, and exports had continued to expand. Broad measures of prices had slowed or contracted in January and February, but excluding energy and food prices, increases in those measures were higher than in previous months. Wage increases had moderated over the past several months.

Example 6. (83% financial market, 17% policy)

Committee members and Board members agreed that, with few exceptions, the functioning of most financial markets, including interbank markets, no longer showed significant impairment. Accordingly they agreed that the statement to be released following the meeting would indicate that the Federal Reserve would be closing the Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility, the Commercial Paper Funding Facility, the Primary Dealer Credit Facility, and the Term Securities Lending Facility on February 1, 2010. Committee members also agreed to announce that temporary liquidity swap arrangements between the Federal Reserve and other central banks would expire on February 1. In addition, the statement would say that amounts available through the Term Auction Facility would be scaled back further, with $50 billion of 28-day credit to be offered on February 8 and $25 billion of 28-day credit to be offered at the final auction of March 8. The statement also would note that the anticipated expiration dates for the Term Asset-Backed Securities Loan Facility remained June 30, 2010, for loans backed by new-issue commercial mortgage-backed securities, and March 31, 2010, for loans backed by all other types of collateral. Members emphasized that they were prepared to modify these plans if necessary to support financial stability and economic growth.

Example 7. (34% growth, 31% financial market, 35% policy)

Open market operations during the intermeeting period continued to be directed toward maintaining the existing degree of pressure on reserve positions. The federal funds rate rose briefly in response to year-end pressures, but it otherwise tended to remain close to the 5-1/4 percent level expected with an unchanged policy stance. Other short-term interest rates generally were unchanged to slightly higher over the intermeeting period. Rates on intermediate- and long-term securities edged higher on balance in reaction to incoming data on economic activity that were on the firm side of market expectations; the increases in such rates appeared to be tempered, however, by favorable market reactions to new data on wages and prices. The generally positive news on economic growth and inflation along with favorable reports on earnings appeared to reinforce the optimism of equity market investors, and major indexes of stock prices increased markedly further over the intermeeting period.

Example 8. (39% growth, 13% inflation, 20% financial market, 26% policy)

In their discussion of the economic situation and outlook, FOMC meeting participants indicated that the worsening financial situation, the slowdown in growth abroad, and incoming information on economic activity had led them to mark down significantly their outlook for growth. While economic activity had evidently already been slowing over the summer, the turmoil in recent weeks had apparently resulted in tighter financial conditions and greater uncertainty among businesses and households about economic prospects, further limiting their ability and willingness to make significant spending commitments. Recent measures of business and consumer sentiment had fallen to historical lows. Participants generally expected the economy to contract moderately in the second half of 2008 and the first half of 2009, and agreed that the downside risks to growth had increased. While some expected an improving financial situation to contribute to a recovery in growth by

mid-2009, others judged that the period of economic weakness could persist for some time. Several participants indicated that they expected some fiscal stimulus in coming quarters, but they were uncertain about the extent and duration of the resulting support to economic activity. Participants agreed that in coming quarters inflation was likely to move down to levels consistent with price stability, reflecting the recent declines in the prices of energy and other commodities, the appreciation of the dollar, and the expected widening of margins of resource slack. Indeed, some saw a risk that over time inflation could fall below levels consistent with the Federal Reserve's dual objectives of price stability and maximum employment.

Figure 1. FOMC Topic Proportions Over Time

This figure plots the proportion of the FOMC minutes devoted to each topic. We use the Latent Dirichlet Allocation (LDA) algorithm to identify the topics. The sample is comprised of 5,644 paragraphs from FOMC minutes released between 1990 and 2015.

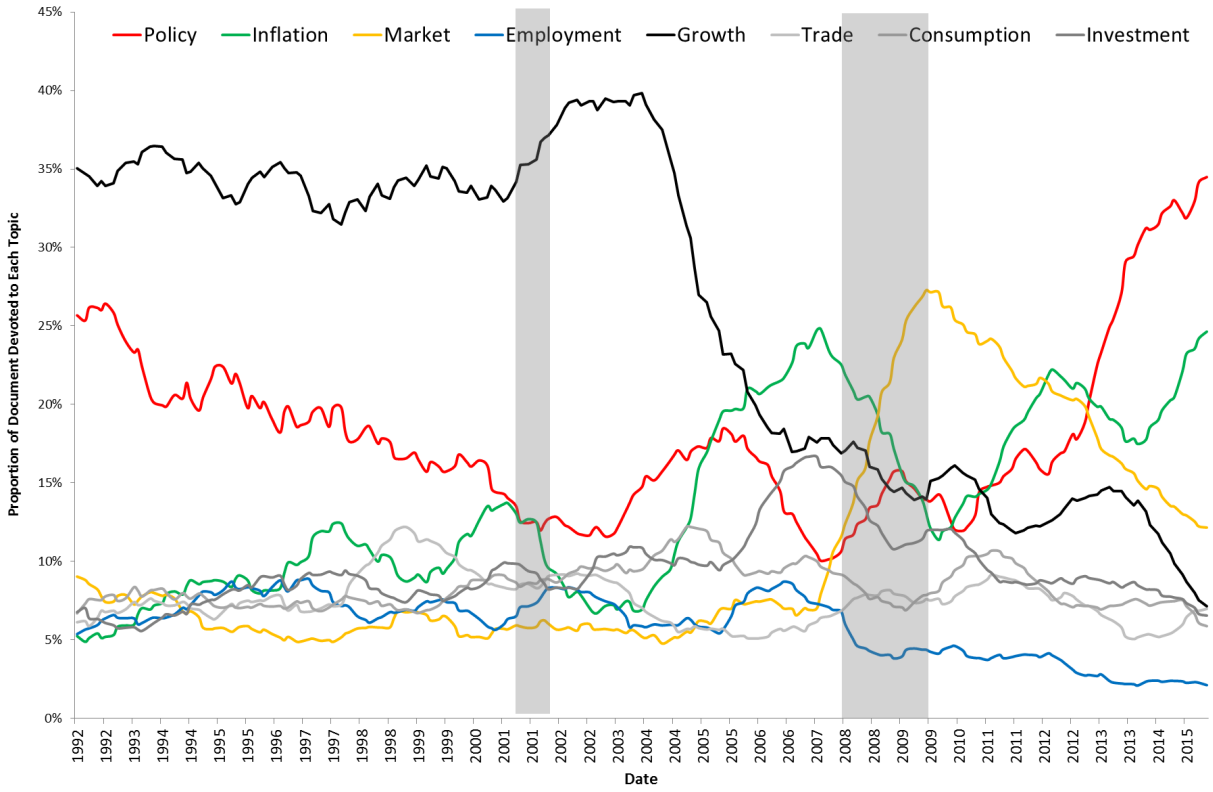
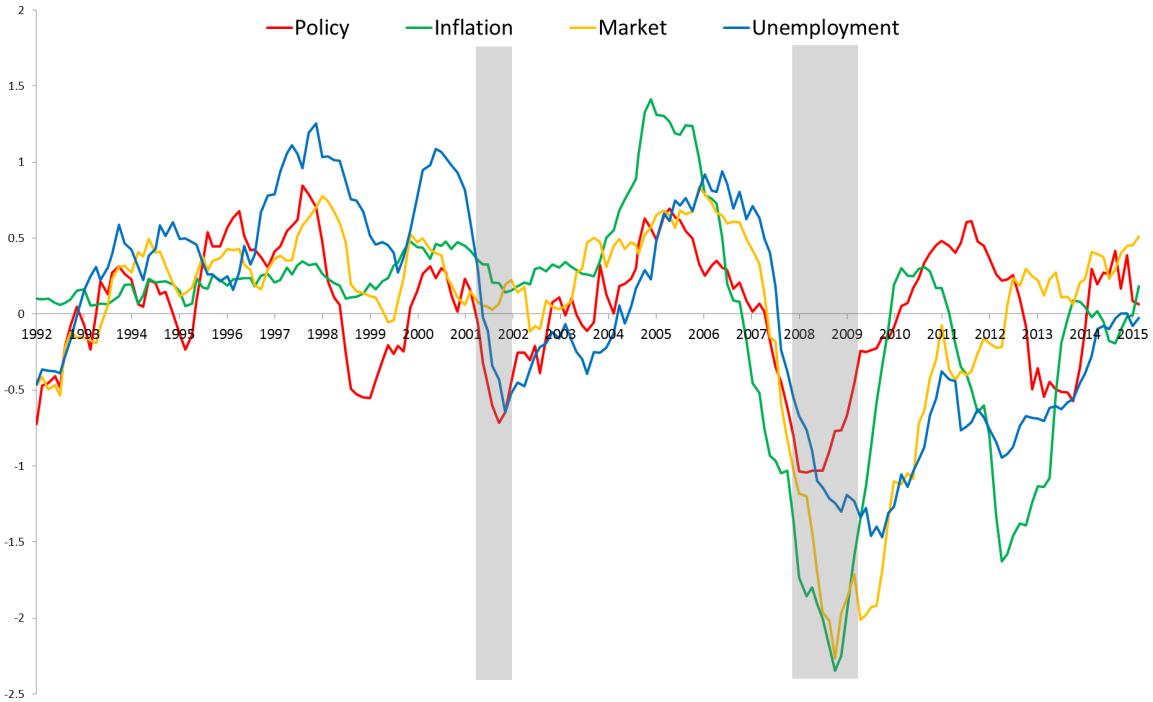
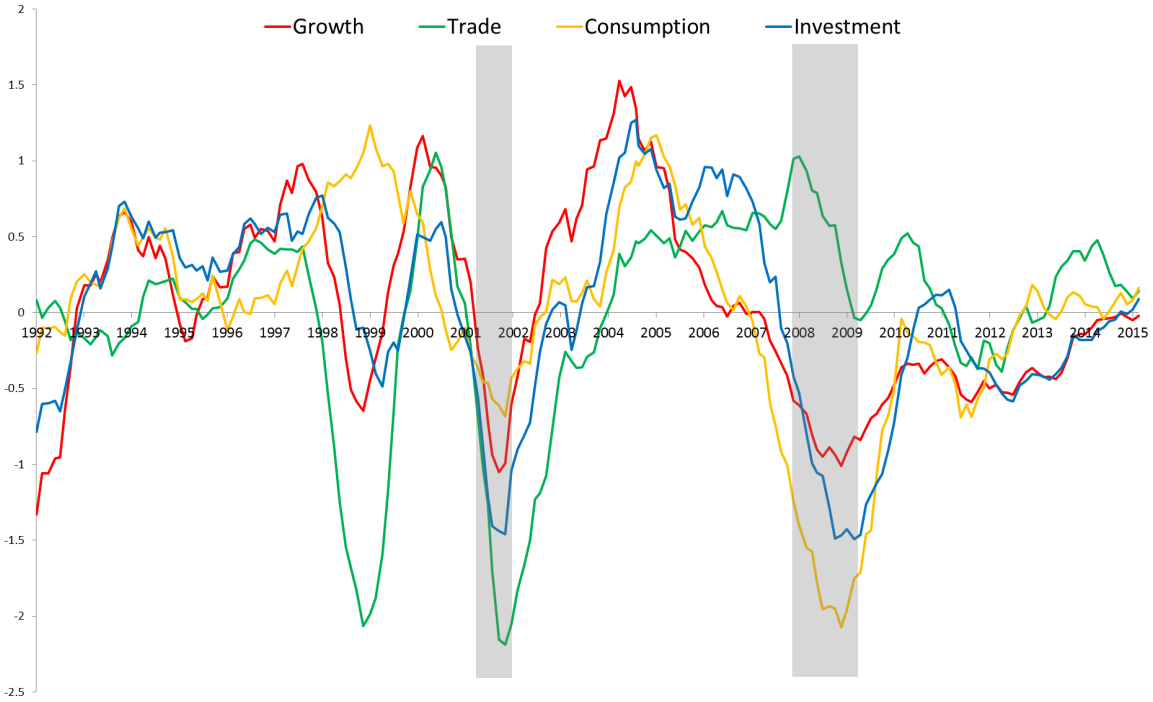


Figure 2. FOMC Topic Content Scores Over Time

This figure plots the tone of each topic discussed in the FOMC minutes. We use the Latent Dirichlet Allocation (LDA) algorithm to identify the topics. The sample is comprised of 5,644 paragraphs from FOMC minutes released between 1990 and 2015.



Topic Tone Scores Over Time, Topics 1 to 4



Topic Tone Scores Over Time, Topics 5 to 8

Figure 3. Volatility Around the Release Time of FOMC Minutes

This figure plots the ratio of average return volatility in 15-minute bins between release and non-release days. Return volatility is calculated as the standard deviation of minute-by-minute returns in each 15-minute bins according to Equation (1) of the text. The solid black line represents the ratio for SPY and the dashed black line represents the ratio for Eurodollars. The sample period is 2000-2015.

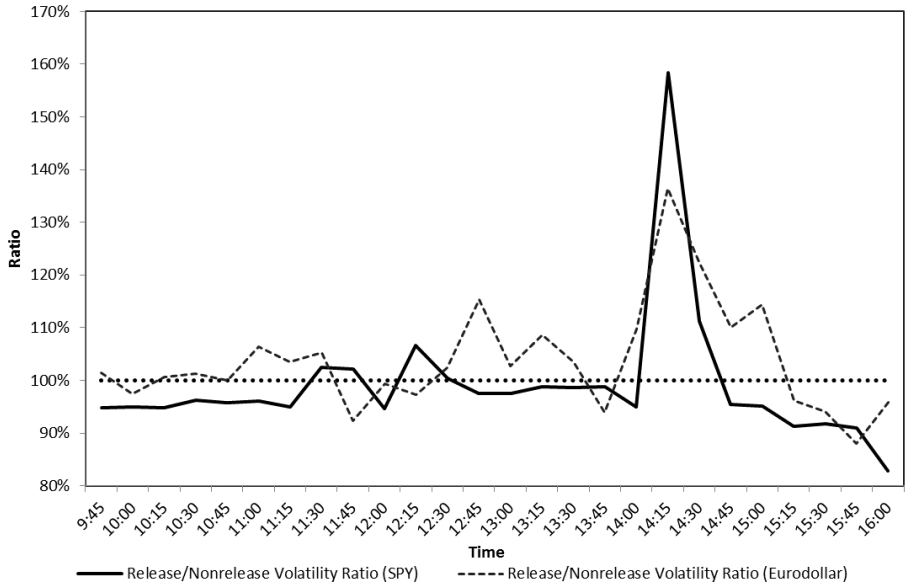
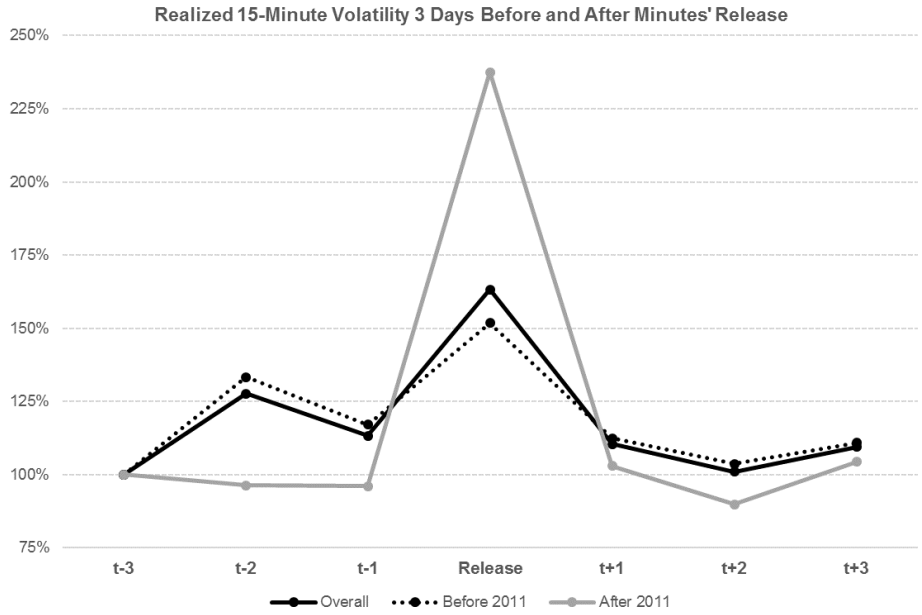


Figure 4. Market Reaction to the Release of FOMC Minutes, Before and After 2011

This figure plots the average realized 15-minute return volatility of the SPY and Eurodollar, from the same window of 2:00pm to 2:15pm, in various subsamples, from $t-3$ to $t+3$ days around the minutes release days t. The volatilities are computed according to Equation (1) of the text. The pre-2011 sample uses 101 FOMC minutes released between 2000 and 2011 for the SPY sample and 60 minutes for the LIBOR sample. The post-2011 sample uses 101 FOMC minutes released between 2011 and 2015 for the SPY sample and 28 minutes for the LIBOR sample. The volatility levels are expressed as a percentage of date $t-3$ volatility.



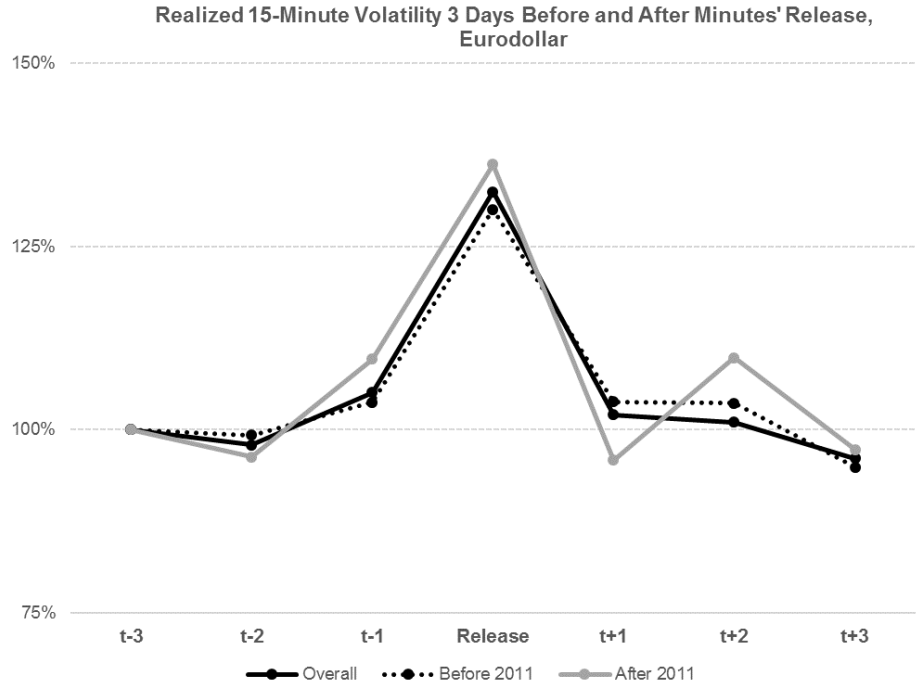
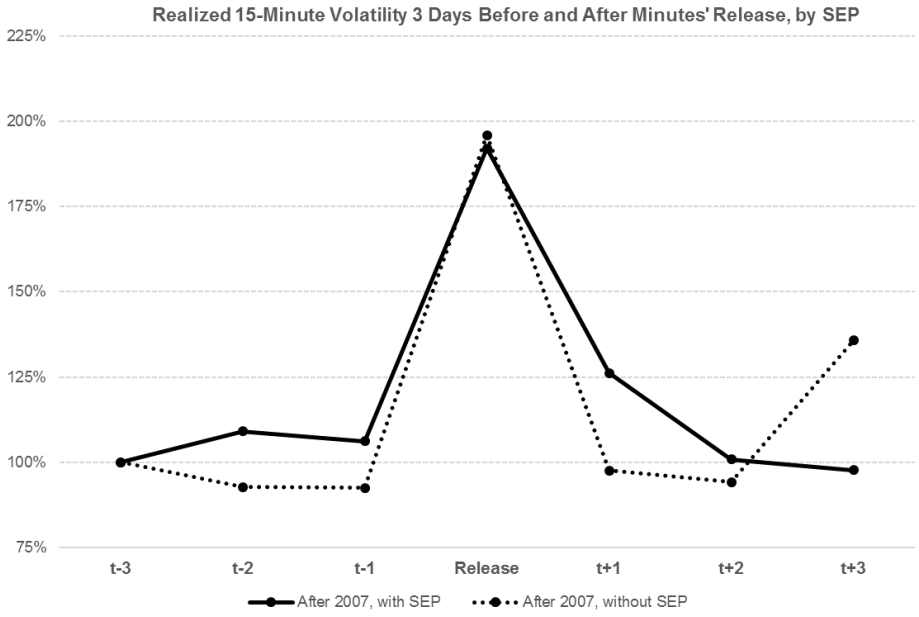


Figure 5. Market Reaction to the Release of FOMC Minutes, With and Without the SEP

This figure plots the average realized 15-minute return volatility of the SPY and Eurodollar, from the same window of 2:00pm to 2:15pm, in various subsamples, from $t-3$ to $t+3$ days around the minutes release days t. The volatilities are computed according to Equation (1) of the text. We plot the raw volatility levels in the post-2007 subsample, between meetings with and without the release of Summary of Economic Projections (SEP) data. The volatility levels are expressed as a percentage of date $t-3$ volatility.



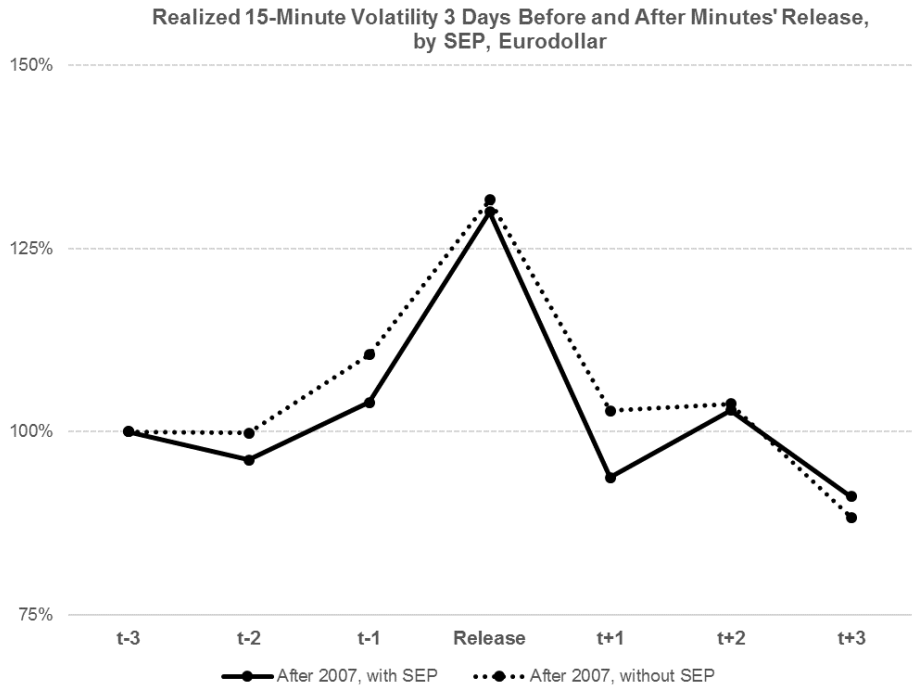


Table I. Distribution of Top LDA Topic Keywords

This table reports the top 20 words for each topic identified by the LDA procedure. Each column in this table represents a topic $k = 1, \ldots, 8$, and the weights are estimates of $\hat{\beta}_{k,j}$ and represent the probability that the word j is generated by topic k. The distributional assumptions for the LDA model are outlined in Section II of the text. The estimation uses 5,644 paragraphs from FOMC meeting minutes released between 1990 and 2015.

Topic 1		Topic 2		Topic 3		Topic 4	
Weight	Word	Weight	Word	Weight	Word	Weight	Word
0.0445	policy	0.0788	inflation	0.0240	market	0.0335	labor
0.0216	monetary	0.0265	energy	0.0206	credit	0.0295	employment
0.0188	funds	0.0255	consumer	0.0172	yields	0.0250	job
0.0143	reserve	0.0226	labor	0.0146	financial	0.0247	workers
0.0133	risks	0.0212	core	0.0144	liquidity	0.0231	payroll
0.0113	financial	0.0178	expectations	0.0142	loans	0.0157	manufacturing
0.0104	agreed	0.0119	compensation	0.0141	securities	0.0151	hiring
0.0100	directive	0.0111	pce	0.0126	debt	0.0147	nonfarm
0.0086	guidance	0.0108	food	0.0123	spreads	0.0138	private
0.0080	purchases	0.0103	unemployment	0.0112	equity	0.0116	unemployment
0.0074	target	0.0099	real	0.0109	corporate	0.0108	inflation
0.0071	stability	0.0090	costs	0.0107	funds	0.0104	hourly
0.0071	easing	0.0089	index	0.0106	commercial	0.0103	services
0.0068	consistent	0.0085	commodity	0.0098	bank	0.0101	earnings
0.0065	stance	0.0082	oil	0.0086	nonfinancial	0.0099	food
0.0063	expectations	0.0072	slack	0.0078	investors	0.0095	costs
0.0057	tightening	0.0069	producer	0.0077	institutions	0.0091	force
0.0056	asset	0.0067	reflecting	0.0075	lending	0.0087	output
0.0054	action	0.0065	subdued	0.0072	issuance	0.0085	utilization
0.0052	view	0.0065	headline	0.0071	bonds	0.0085	construction

Topic 5		Topic 6		Topic 7		Topic 8	
Weight	Word	Weight	Word	Weight	Word	Weight	Word
0.0208	economy	0.0340	foreign	0.0448	consumer	0.0447	production
0.0169	business	0.0315	exports	0.0381	sales	0.0369	manufacturing
0.0129	economic	0.0289	u.s	0.0335	housing	0.0354	inventories
0.0111	demand	0.0268	dollar	0.0168	homes	0.0275	output
0.0087	productivity	0.0223	imports	0.0165	mortgage	0.0266	motor
0.0076	investment	0.0219	economies	0.0164	starts	0.0223	investment
0.0072	pressure	0.0166	countries	0.0145	construction	0.0201	industrial
0.0068	firms	0.0152	trade	0.0138	income	0.0160	sales
0.0063	financial	0.0140	major	0.0135	household	0.0149	equipment
0.0058	fiscal	0.0128	currencies	0.0134	gains	0.0143	vehicles
0.0057	prospects	0.0125	industrial	0.0131	expenditures	0.0141	business
0.0056	capital	0.0118	deficit	0.0105	single-family	0.0136	stocks
0.0055	confidence	0.0117	united	0.0101	retail	0.0118	wholesale
0.0055	strength	0.0112	japan	0.0098	motor	0.0118	capacity
0.0054	sectors	0.0098	exchange	0.0097	personal	0.0112	utilization
0.0053	potential	0.0097	euro	0.0091	purchases	0.0097	ratio
0.0051	favorable	0.0088	emerging	0.0078	vehicles	0.0090	industries
0.0050	costs	0.0084	sovereign	0.0077	existing	0.0087	retail
0.0049	anecdotal	0.0080	abroad	0.0076	residential	0.0074	accumulation
0.0049	stimulus	0.0072	european	0.0073	sentiment	0.0072	factory

Table II. Topic Proportion and Macroeconomic Variables

This table reports the coefficient estimates for Regression (5). The dependent variables are document-level proportions for each topic. The independent variables are the following: $IntRate$: the latest daily closing yield of 10-year Treasury notes obtained from the Federal Reserve Board's H.15 releases; $UnEmp$: latest monthly rate of unemployment obtained from the Bureau of Labor Statistics; and $Recession$: a dummy variable which equal to one if the FOMC date falls within a NBER-designated recession period. These variables are also used as controls in other regressions. The estimates use 196 FOMC minutes released between 1990 and 2015.

	Topics							
	(1) Policy	(2) Inflation	(3) Market	(4) Employment	(5) Growth	(6) Trade	(7) Consumption	(8) Investment
Interest Rate	0.0487	-0.432***	-0.221***	-0.107**	0.409***	0.00231	-0.164***	0.244***
	(1.06)	(-13.12)	(-8.96)	(-2.47)	(13.46)	(0.05)	(-3.71)	(6.38)
Unemployment	0.0690	-0.0425	0.345***	-0.156***	-0.0799**	-0.0694	-0.179***	-0.158***
	(1.42)	(-1.22)	(13.23)	(-3.38)	(-2.48)	(-1.42)	(-3.81)	(-3.90)
Recession Dummy	-0.446*	-0.0926	0.974***	-0.913***	-0.102	0.227	0.153	-0.138
	(-1.85)	(-0.54)	(7.56)	(-4.00)	(-0.64)	(0.94)	(0.66)	(-0.69)
N	196	196	196	196	196	196	196	196
adj. R-sq	0.015	0.493	0.717	0.114	0.568	0.001	0.085	0.315

Table III. Topic Tone Scores and Macroeconomic Variables

This table reports the coefficient estimates for Regression (7). The independent variables are Tone Scores for each of the eight LDA-identified topics computed according to Equation (6) of the text. The independent variables are the following: $IntRate$: the latest daily closing yield of 10-year Treasury notes obtained from the Federal Reserve Board's H.15 releases; $UnEmp$: latest monthly rate of unemployment obtained from the Bureau of Labor Statistics; and $Recession$: a dummy variable which equal to one if the FOMC date falls within a NBER-designated recession period. The estimates use 196 FOMC minutes released between 1990 and 2015.

Panel A: Net Tone

	Topic							
	Policy	Inflation	Market	Employment	Economy	Trade	Consumption	Investment
Interest Rate	0.0036	0.1160***	-0.0129	0.1334***	-0.0029	0.0030	0.0475	0.0536
	(0.08)	(2.78)	(-0.31)	(3.66)	(-0.07)	(0.06)	(1.16)	(1.35)
Unemployment	0.0308	-0.0481	-0.1683***	-0.2332***	-0.1398**	0.0294	-0.0759*	-0.0965**
	(0.65)	(-1.09)	(-3.81)	(-6.03)	(-3.15)	(0.60)	(-1.76)	(-2.30)
Recession	-0.9529***	-1.2035***	-1.2021***	-1.1137***	-1.2576***	-0.4359*	-1.4928***	-1.5910***
	(-4.07)	(-5.50)	(-5.50)	(-5.83)	(-5.73)	(-1.80)	(-6.99)	(-7.67)
N	196	196	196	196	196	196	196	196
adj. R-sq	0.067	0.185	0.187	0.378	0.180	0.003	0.222	0.266

Table IV. Market Reaction to the Release of FOMC Minutes

This table reports the coefficient estimates for Regression (8), fitted using the release-day dummy, as well as volatilities over the past 0, 5, 10, and 20 trading days. The number of lagged trading days we use as control is indicated on the top of each column. The equity market regression uses transaction prices of SPY to measure intraday market returns. The LIBOR regression uses 3-month LIBOR implied by the nearest maturity Eurodollar futures contract prices. Each regression uses between 4,343 and 4,363 days of observation.

Equity Market	Number of Lags in Control			
	(0)	(5)	(10)	(20)
Release Dummy	0.5919***	0.6081***	0.6130***	0.6032***
	(6.32)	(6.99)	(7.13)	(7.02)
No. Obs	4363	4358	4353	4343
adj. R-sq	0.011	0.149	0.17	0.182
LIBOR	Number of Lags in Control			
	(0)	(5)	(10)	(20)
Release Dummy	0.3054**	0.2283*	0.2665**	0.2638*
	(2.02)	(1.74)	(2.01)	(1.98)
No. Obs	2627	2622	2617	2607
adj. R-sq	0.004	0.124	0.136	0.151

Table V. Market Reaction to the Overall Content of FOMC Minutes

This table reports the coefficient estimates for Regression (9). The dependent variable is the 15-minute unexpected volatility computed as the raw volatility minus the 20-day moving average, according to Equation 10 in the text. The independent variables are document-level tone scores computed according to Equation (6). The control variables are: $IntRate$: the latest daily closing yield of 10-year Treasury notes obtained from the Federal Reserve Board's H.15 releases; $UnEmp$: latest monthly rate of unemployment obtained from the Bureau of Labor Statistics; and $Recession$: a dummy variable which equal to one if the FOMC date falls within a NBER-designated recession period. The estimates for SPY use 138 FOMC minutes released between 2000 and 2015, and the estimates for LIBOR uses 88 minutes released between 2003 and 2014. The stars (***, **, *) denote statistical significance levels at 1%, 5%, and 10%, respectively.

	Document-Level Net Tone	
	SPY Data	LIBOR Data
	(1)	(2)
Document Tone	-0.0159*	0.0037
	(-1.69)	(0.78)
Interest Rate	-0.007	-0.0096
	(-0.82)	(-1.20)
Unemployment	-0.0062	-0.0080
	(-1.03)	(-1.35)
Recession	-0.0486*	0.0093
	(-1.75)	(1.42)
N	138	88
R-sq	0.008	0.017

Table VI. FOMC Topic Proportion and Market Reaction

Columns (1) to (4) of this table report the coefficient estimates for Regression (11a). The dependent variable is the 15-minute unexpected volatility computed as the raw volatility minus the 20-day moving average of volatility, according to Equation (10) in the text. Columns (5) to (8) of this table report the coefficient estimates where the dependent variable is directional change in SPY or LIBOR. The regressions use transaction prices of SPY and 3-month LIBOR implied by the nearest maturity Eurodollar futures contract prices to compute intraday unexpected volatility and Directional Price Changes. The independent variables are document-level proportions for each topic. The control variables are: $IntRate$: the latest daily closing yield of 10-year Treasury notes obtained from the Federal Reserve Board's H.15 releases; $UnEmp$: latest monthly rate of unemployment obtained from the Bureau of Labor Statistics; and $Recession$: a dummy variable which equal to one if the FOMC date falls within a NBER-designated recession period. The estimates for SPY use 138 FOMC minutes released between 2000 and 2015, and the estimates for LIBOR use 88 minutes released between 2003 and 2014. The stars (***, **, *) denote statistical significance levels at 1%, 5%, and 10%, respectively.

	Unexpected Volatility				Directional Price Change			
	SPY		LIBOR		SPY		LIBOR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Policy	0.0193**	0.0198**	0.2118**	0.0256***	0.0398	0.0640*	-0.2083*	-0.1218*
	(2.24)	(2.49)	(2.51)	(2.69)	(1.38)	(1.71)	(-1.94)	(-1.80)
Inflation	0.0208**	0.0201**	0.1712*	0.1939**	-0.0079	0.0125	-0.1010	-0.1679
	(2.09)	(2.46)	(1.89)	(2.20)	(-0.57)	(1.12)	(-0.74)	(-0.20)
Market	-0.0007	-0.0219	0.0303	-0.0146	-0.0209	-0.0245	0.0954	0.1384
	(-0.12)	(-1.21)	(0.26)	(-0.21)	(-0.74)	(-0.90)	(0.86)	(1.02)
Employment	0.0153*	0.0144**	0.2044**	0.2953**	0.0371	0.0265	-0.1472*	-0.1893
	(1.94)	(2.42)	(2.50)	(2.35)	(1.47)	(1.08)	(-1.73)	(-1.50)
Growth	0.0039	0.0046	0.1432	0.1027	0.0172	0.0093	-0.0001	-0.0043
	(1.52)	(1.28)	(0.98)	(1.07)	(0.68)	(0.24)	(-0.02)	(-0.20)
Trade	-0.0032	0.0008	-0.0729	-0.1256	-0.0334	-0.0146	0.1454	0.1110
	(-0.27)	(0.37)	(-0.15)	(-0.65)	(-0.30)	(-0.22)	(1.38)	(1.13)
Consumption	-0.0017	-0.0074	-0.1175*	-0.1453	-0.0013	0.0030	-0.0079	-0.0235
	(-0.75)	(-1.13)	(-1.77)	(-1.56)	(-0.71)	(0.22)	(-0.85)	(-1.07)
Investment	0.0287*	0.0265**	-0.0240	-0.0134	0.0301*	0.0238	-0.0034	-0.0062
	(1.94)	(2.49)	(-0.60)	(-0.84)	(1.69)	(1.32)	(-0.11)	(-0.08)
Control Variables								
Interest Rate		-0.0011		0.0269		0.0124		0.1441
		(-0.16)		(1.23)		(0.65)		(1.20)
Unemployment		0.0214*		0.1793		0.0201*		0.0788
		(1.88)		(1.47)		(1.74)		(0.32)
Recession		0.0120		0.4340		-0.1104		-0.0105
		(0.53)		(1.39)		(-0.95)		(-0.33)
N	138	138	88	88	138	138	88	88
adj. R-sq	0.092	0.109	0.074	0.080	0.001	0.016	0.015	0.019

Table VII. FOMC Topic Tone Score and Market Reaction

Columns (1) to (4) of this table report the coefficient estimates for Regression (11b). The dependent variable is the 15-minute unexpected volatility computed as the raw volatility minus the 20-day moving average, according to Equation 10 in the text. Columns (5) to (8) of this table report the coefficient estimates where the dependent variable is directional change in SPY or LIBOR. The regressions use transaction prices of SPY and 3-month LIBOR implied by the nearest maturity Eurodollar futures contract to compute intraday unexpected volatility and Directional Price Changes. The independent variables are document-level proportions for each topic. The control variables are: $IntRate$: the latest daily closing yield of 10-year Treasury notes obtained from the Federal Reserve Board's H.15 releases; $UnEmp$: latest monthly rate of unemployment obtained from the Bureau of Labor Statistics; and $Recession$: a dummy variable which equal to one if the FOMC date falls within a NBER-designated recession period. The independent variables are document-level Net Tone scores for each of the eight LDA-identified topics, computed according to Equation (6). The estimates for SPY use 138 FOMC minutes released between 2000 and 2015, and the estimates for LIBOR use 88 minutes released between 2003 and 2014. The stars (***, **, *) denote statistical significance levels at 1%, 5%, and 10%, respectively.

	Unexpected Volatility				Directional Price Change			
	SPY		LIBOR		SPY		LIBOR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Policy	-0.0014	-0.0137	-0.3448**	-0.3512**	0.0609**	0.0670***	-0.4146***	-0.4412**
	(-0.20)	(-1.48)	(-2.58)	(-2.23)	(2.44)	(2.70)	(-2.62)	(-2.38)
Inflation	-0.0180**	-0.0187***	-0.3719***	-0.3723***	0.0443***	0.0445**	-0.4740**	-0.4839***
	(-2.34)	(-2.77)	(-3.11)	(-3.25)	(2.76)	(2.49)	(-2.58)	(-2.67)
Market	-0.0047	-0.0024	0.0656	0.1001	-0.0560***	-0.0556***	0.1018	0.0693
	(-0.69)	(-0.32)	(0.20)	(0.30)	(-2.73)	(-2.65)	(0.74)	(0.62)
Employment	-0.0155**	-0.0154**	-0.1731**	-0.2087*	0.0433*	0.0601**	-0.0994*	-0.2330
	(-2.46)	(-2.33)	(-2.04)	(-1.86)	(1.91)	(2.36)	(-1.87)	(-1.59)
Growth	0.0279*	0.0350**	0.1753	0.1386	0.0025	-0.0143	-0.1439	-0.1171
	(1.88)	(2.26)	(1.54)	(1.49)	(0.14)	(-0.82)	(-0.65)	(-1.37)
Trade	-0.0075	-0.0096	-0.4511	-0.0797	0.0030	0.0050	0.0097	0.0126
	(-0.84)	(-1.05)	(-0.12)	(-0.04)	(0.16)	(0.03)	(0.38)	(0.35)
Consumption	0.0157**	0.0144**	0.1695	0.0878	-0.0364	-0.0539**	0.3381	0.1917
	(2.00)	(2.02)	(1.43)	(1.19)	(-1.38)	(-2.03)	(0.72)	(0.52)
Investment	-0.0131	-0.0136	0.0901	-0.0478	-0.0595**	-0.0698***	-0.0015	-0.0123
	(-1.35)	(-1.34)	(0.07)	(-0.82)	(-2.55)	(-3.00)	(-0.81)	(-0.86)
Control Variables								
Interest Rate		-0.0002		0.1683		0.0195		0.1200
		(-0.98)		(0.64)		(1.07)		(0.52)
Unemployment		0.0221*		0.0474		0.0169		-0.1477
		(1.69)		(1.25)		(1.05)		(-0.21)
Recession		0.0173*		-0.0735		-0.1170*		-0.0093
		(1.70)		(-0.14)		(-1.84)		(-0.36)
N	138	138	88	88	138	138	88	88
adj. R-sq	0.069	0.077	0.084	0.082	0.015	0.025	0.012	0.009

Table VIII. Market Reaction to Minutes Released Before and After 2011

This table reproduces Table VII for different subsamples. The regressions use transaction prices of SPY and 3-month LIBOR implied by the nearest maturity Eurodollar futures contract prices to compute intraday unexpected volatility and Directional Price Changes. The independent variables are document-level proportions for each topic. The control variables are: $IntRate$: the latest daily closing yield of 10-year Treasury notes obtained from the Federal Reserve Board's H.15 releases; $UnEmp$: latest monthly rate of unemployment obtained from the Bureau of Labor Statistics; and $Recession$: a dummy variable which equal to one if the FOMC date falls within a NBER-designated recession period. The pre-2011 sample uses 101 FOMC minutes released between 2000 and 2011 for the SPY sample and 60 minutes for the LIBOR sample. The post-2011 sample uses 101 FOMC minutes released between 2011 and 2015 for the SPY sample and 28 minutes for the LIBOR sample. The estimates for SPY use 101 FOMC minutes released between 2000 and 2010, and 37 FOMC minutes between 2011 and 2015. The estimates for LIBOR use 60 FOMC minutes released between 2003 and 2010, and 28 FOMC minutes released between 2010 and 2015. The stars (***, **, *) denote statistical significance levels at 1%, 5%, and 10%, respectively.

Topic	Unexpected Volatility				Directional Price Change			
	SPY		LIBOR		SPY		LIBOR	
	Pre-2011	Post-2011	Pre-2011	Post-2011	Pre-2011	Post-2011	Pre-2011	Post-2011
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Policy	-0.0127	-0.0243	-0.2829*	-0.4103*	0.0649***	0.0711**	-0.5101**	-0.3574*
	(-1.31)	(-1.57)	(-1.93)	(-1.71)	(-2.69)	(-2.53)	(-2.50)	(-1.96)
Inflation	-0.0128**	-0.0219**	-0.3154***	-0.3970	0.0737**	0.0203*	-0.4537***	-0.5061**
	(-2.37)	(-2.47)	(-3.37)	(-1.55)	(-2.32)	(-1.72)	(-2.76)	(-2.09)
Market	-0.0043	-0.005	-0.1102	0.0023	-0.0526**	-0.0741***	0.0820	-0.0372*
	(-1.09)	(-0.37)	(-0.56)	(0.20)	(-1.99)	(-3.65)	(0.44)	(-1.81)
Employment	-0.012**	-0.0279**	-0.1955*	-0.2402	0.0632**	0.0642**	-0.1872**	-0.3150
	(-2.22)	(-2.42)	(-1.84)	(-1.35)	(-2.20)	(-2.65)	(-2.13)	(-0.64)
Growth	-0.0162*	0.0297**	0.1413*	0.0574	0.0202	-0.0088	-0.1352	-0.1026
	(-1.98)	(-2.34)	(1.70)	(0.39)	(-0.32)	(-0.72)	(-0.91)	(-1.58)
Trade	-0.0089	-0.0018	-0.0981	0.0657	0.0127	0.0028	0.0230	-0.0625
	(-1.31)	(-0.18)	(-1.02)	(0.81)	(-0.54)	(-0.12)	(0.79)	(-0.42)
Consumption	0.0308***	0.0153	0.0640	0.1125	-0.05	-0.0442	0.1828	0.4555
	(-3.01)	(-1.14)	(1.38)	(1.22)	(-0.13)	(-0.24)	(1.14)	(0.06)
Investment	0.0024	-0.0351**	0.0104	-0.1342	-0.077**	-0.0628**	0.0224	-0.2391
	(-0.26)	(-2.34)	(1.13)	(-0.15)	(-2.60)	(-2.14)	(0.90)	(-1.13)
Control Variables								
Interest Rate	-0.0022	0.0016	0.1192	0.3213	0.0071	0.1572**	0.1403	0.0068
	(-0.13)	(-0.54)	(1.39)	(0.40)	(-0.68)	(-2.31)	(0.95)	(0.14)
Unemployment	0.0117	0.0114	0.0460	0.1871	0.0105	0.0155	-0.3851	0.0513
	(1.62)	(-0.17)	(1.45)	(0.78)	(-0.58)	(-0.72)	(-0.47)	(0.48)
Recession	0.0179		0.0251		-0.0463		-0.0118	
	(0.71)		(0.99)		(-0.78)		(-0.62)	
N	101	37	60	28	101	37	60	28
adj. R-sq	0.04	0.089	0.071	0.093	0.018	0.172	0.008	0.024

Table IX. Market Reaction to Minutes With and Without Concurrent SEP Release

This table replicates Table VII for the subsamples with and without concurrent Summary of Economic Projections (SEP) releases. The first SEP is released for the October 2007 meeting, and our overall sample is therefore all FOMC minutes released after June 2007, consisting of 64 minutes. Of these minutes, 31 are with concurrent SEP releases and 33 are without. We reproduce the tone regressions using both the 15-minute unexpected volatility and directional price change of the SPY and LIBOR implied by the nearest maturity Eurodollar futures contract price. The control variables are: $IntRate$: the latest daily closing yield of 10-year Treasury notes obtained from the Federal Reserve Board's H.15 releases; $UnEmp$: latest monthly rate of unemployment obtained from the Bureau of Labor Statistics; and $Recession$: a dummy variable which equal to one if the FOMC date falls within a NBER-designated recession period. The SPY regression estimates use 32 FOMC minutes released between 2007 and 2015. The LIBOR regression estimates use 27 FOMC minutes with the SEP and 29 FOMC minutes without the SEP, released between 2007 and 2014.

Panel A: SPY Regressions

Topic	Unexpected Volatility		Directional Price Change	
	With SEP	Without SEP	With SEP	Without SEP
Policy	-0.0119	-0.0108*	0.0797*	0.0714**
	(-1.53)	(-1.89)	(1.97)	(2.35)
Inflation	-0.0202**	-0.0330**	0.0180**	0.0393**
	(-2.12)	(-2.30)	(2.64)	(2.72)
Market	-0.0045	-0.0263	-0.0387**	-0.0671
	(-0.29)	(-1.47)	(-2.11)	(-1.60)
Employment	-0.0133	-0.0192*	0.0304	0.0581
	(-1.49)	(-1.73)	(1.38)	(0.95)
Growth	0.0384*	0.0573	-0.0198	-0.0262*
	(1.74)	(1.48)	(-1.40)	(-1.77)
Trade	0.0038	-0.0159	-0.0732	0.0008
	(0.47)	(-0.60)	(-0.74)	(0.05)
Consumption	0.0075	0.0192**	-0.0374	-0.0523
	(0.61)	(2.14)	(-0.99)	(-1.68)
Investment	-0.0102	0.0258	-0.0545**	-0.0312***
	(-0.48)	(0.31)	(-2.10)	(-2.79)
Control Variables	Yes	Yes	Yes	Yes
N	32	32	32	32
adj. R-sq	0.140	0.153	0.059	0.075

Panel B: LIBOR Regressions

Topic	Unexpected Volatility		Directional Price Change	
	With SEP	Without SEP	With SEP	Without SEP
Policy	-0.3338*	-0.4200**	-0.3819*	-0.4262**
	(-1.90)	(-2.34)	(-1.89)	(-2.61)
Inflation	-0.1195	-0.3002*	-0.5477*	-0.4702
	(-1.27)	(-1.74)	(-1.80)	(-1.65)
Market	0.0482	-0.1738	-0.0043	-0.0388
	(0.28)	(-1.41)	(-0.62)	(-0.44)
Employment	-0.2337**	-0.2702	-0.1404*	-0.1736*
	(-2.14)	(-1.26)	(-2.03)	(-1.92)
Growth	0.1506*	0.1064	-0.1447	0.0019
	(1.84)	(0.82)	(-0.53)	(0.87)
Trade	0.0017	-0.1179	0.0426	-0.0081
	(0.21)	(-0.28)	(1.24)	(-0.31)
Consumption	0.0607	0.1433	0.3853	0.2396
	(0.70)	(0.57)	(1.56)	(1.39)
Investment	-0.0014	-0.1287	0.0255	-0.1905
	(-0.04)	(-0.49)	(0.47)	(-0.12)
Control Variables	Yes	Yes	Yes	Yes
N	27	29	27	29
adj. R-sq	0.113	0.046	0.02	0.007

Table X. Market Reaction to Tone Computed Using Alternative Lexicons and Weighting Schemes

This table replicates Table VII for different tone scores computed using alternative tonal lexicons and weighting schemes. Columns (1)-(2) and (5)-(6) of this table reproduce the unexpected volatility and directional price change regressions using the Harvard IV-4 and Loughran and McDonald (2011) lexicons, respectively. Columns (3) and (7) use the combined lexicon, but replaces the equal term weights with tf.idf term weights. Finally, Columns (4) and (8) replace the equal term weights with weights derived from market reactions to 10-Ks developed by Jegadeesh and Wu (2013). The independent variables are document-level tone scores for each of the eight LDA-identified topics, computed according to Equation (6), using these alternative lexicons and weighting schemes. The control variables are: $IntRate$: the latest daily closing yield of 10-year Treasury notes obtained from the Federal Reserve Board's H.15 releases; $UnEmp$: latest monthly rate of unemployment obtained from the Bureau of Labor Statistics; and $Recession$: a dummy variable which equal to one if the FOMC date falls within a NBER-designated recession period. The estimates use 138 FOMC minutes released between 2000 and 2015, and the estimates for LIBOR use 88 minutes released between 2003 and 2014.

Panel A: SPY Regressions

Topic	Unexpected Volatility				Directional Price Change			
	Lexicon		Weighting		Lexicon		Weighting	
	IV4	LM	tf.idf	WP	IV4	LM	tf.idf	WP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Policy	-0.0156	-0.0194	-0.0292	-0.0073**	0.0659***	0.0578***	0.0742**	0.0544*
	(-1.55)	(-1.12)	(-1.06)	(-2.11)	(2.63)	(2.85)	(2.58)	(1.84)
Inflation	-0.0154***	-0.0220**	-0.0169**	-0.0135**	0.0505**	0.0492**	0.0403***	0.0133
	(-2.80)	(-2.51)	(-2.19)	(-2.54)	(2.55)	(2.09)	(2.71)	(0.60)
Market	-0.0073	-0.0132	-0.0115	-0.0127*	-0.0493**	-0.0564***	-0.0318	-0.0474***
	(-0.45)	(-0.28)	(-1.53)	(-1.96)	(-2.21)	(-2.97)	(-1.33)	(-2.94)
Employment	-0.0126***	-0.0214**	-0.0181*	-0.0019	0.0635**	0.0571**	0.0392*	0.0597**
	(-2.71)	(-1.99)	(-1.94)	(-0.36)	(2.47)	(2.01)	(1.94)	(2.12)
Growth	0.0298**	0.0370**	0.0353**	0.0188***	-0.0172	-0.0126	-0.0198	0.0754***
	(2.40)	(2.04)	(2.29)	(3.13)	(-1.24)	(-1.28)	(-1.13)	(3.10)
Trade	-0.0072	0.0008	-0.0026	-0.0142	0.0094	0.0065	-0.0005	0.0519
	(-0.84)	(0.12)	(-0.30)	(-1.26)	(0.28)	(0.50)	(-0.19)	(1.45)
Consumption	0.0153*	0.0084	0.0179**	-0.024***	-0.0385	-0.0662**	-0.0428*	-0.0086
	(1.95)	(0.66)	(2.01)	(-2.94)	(-1.06)	(-2.49)	(-1.87)	(-0.33)
Investment	-0.0117	0.0075	-0.0155	0.0343***	-0.0602**	-0.0624*	-0.0712**	-0.0709**
	(-1.40)	(1.59)	(-1.06)	(3.92)	(-2.20)	(-1.87)	(-2.11)	(-2.25)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	138	138	138	138	138	138	138	138
adj. R-sq	0.082	0.069	0.060	0.077	0.014	0.027	0.015	0.033

Panel B: LIBOR Regressions

Policy	-0.3004**	-0.2891*	-0.2652***	-0.3943***	-0.3680*	-0.4662**	-0.4332**	-0.4071*
	(-2.57)	(-1.74)	(-2.78)	(-2.71)	(-1.82)	(-2.55)	(-2.07)	(-1.93)
Inflation	-0.3934**	-0.3439***	-0.3949***	-0.3614*	-0.4807*	-0.5375*	-0.4781***	-0.5065***
	(-2.30)	(-3.32)	(-3.37)	(-2.00)	(-1.95)	(-1.79)	(-2.84)	(-2.74)
Market	-0.1587	-0.1733	0.0338	-0.1102	0.0081	0.0733	-0.0271	0.0513
	(-0.96)	(-1.50)	(0.72)	(-0.41)	(0.31)	(1.00)	(-0.08)	(0.65)
Employment	-0.2233**	-0.1714*	-0.1003	-0.3012**	-0.1113	-0.2158*	-0.1794	-0.1518
	(-2.42)	(-1.70)	(-1.56)	(-2.48)	(-0.79)	(-1.83)	(-1.28)	(-0.56)
Growth	0.1038	0.0822	0.2433	0.1775	-0.0823	0.0012	-0.0979	-0.0138
	(0.87)	(1.35)	(0.72)	(0.42)	(-0.00)	(0.11)	(-1.44)	(-1.35)
Trade	0.0072	-0.0237	-0.0740	0.0010	-0.0107	0.0550	0.0824	-0.0442
	(0.83)	(-0.19)	(-0.47)	(0.06)	(-0.38)	(0.74)	(0.40)	(-0.15)
Consumption	0.0692	0.1039	0.0114	0.1507	0.2180	-0.0912	0.2554	0.0846
	(1.17)	(0.07)	(0.99)	(1.42)	(1.10)	(-0.29)	(1.17)	(0.90)
Investment	-0.0581	0.0291	-0.0365	-0.1013	-0.0935	0.1002	0.0297	0.1158
	(-0.21)	(0.87)	(-0.38)	(-0.55)	(-0.02)	(0.37)	(0.91)	(0.75)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	88	88	88	88	88	88	88	88
adj. R-sq	0.072	0.094	0.088	0.069	0.007	0.013	0.019	0.014