NBER WORKING PAPER SERIES

MEASURING CENTRAL BANK COMMUNICATION: AN AUTOMATED APPROACH WITH APPLICATION TO FOMC STATEMENTS

David O. Lucca
Francesco Trebbi

Working Paper 15367
http://www.nber.org/papers/w15367

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2009

This version of the paper extends the analysis of the draft dated April, 2008 in constructing linguistic measures based on documents from the DJ Factiva news database, in addition to the Internet-based ones. The authors would like to thank Matilde Bombardini, John Cochrane, Erik Hurst, Anil Kashyap, Monika Piazzesi, Jonathan Wright and Egon Zakrajsek for their comments and suggestions, as well as participants at various seminars and Ken Kuttner, our discussant at the meeting of the NBER Monetary Economics Program. Isaac Laughlin provided outstanding research assistance. Hal Varian and the Google University Research Program provided generous access to the Google Search technology. The views and analysis set forth are solely those of the authors and do not indicate concurrence by other members of the Board of Governors of the Federal Reserve System. Francesco Trebbi gratefully acknowledges support by the Initiative on Global Markets and the Mutch Family Faculty Research Fund at the Booth School of Business, the University of Chicago. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2009 by David O. Lucca and Francesco Trebbi. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements
David O. Lucca and Francesco Trebbi
NBER Working Paper No. 15367
September 2009
JEL No. E43,E52,E58

ABSTRACT

We present a new automated, objective and intuitive scoring technique to measure the content of central bank communication about future interest rate decisions based on information from the Internet and news sources. We apply the methodology to statements released by the Federal Open Market Committee (FOMC) after its policy meetings starting in 1999. Using intra-day financial quotes, we find that short-term nominal Treasury yields respond to changes in policy rates around policy announcements, whereas longer-dated Treasuries mainly react to changes in policy communication. Using lower frequency data, we find that changes in the content of the statements lead policy rate decisions by more than a year in univariate interest rate forecasting and vector autoregression (VAR) models. When we estimate Treasury yield responses to the shocks identified in the VAR, we find communication to be a more important determinant of Treasury rates than contemporaneous policy rate decisions. These results are consistent with the view that the FOMC releases information about future policy rate actions in its statements and that market participants incorporate this information when pricing longer-dated Treasuries. Finally, we decompose realized policy rate decisions using a forward-looking Taylor rule model. Based on this decomposition, we find that FOMC statements contain significant information regarding both the predicted rule-based interest rate and the Taylor-rule residual component, and that content of the statements leads the residual by a few quarters.

David O. Lucca
Mailstop 84
Board of Governors of the Federal Reserve System
Washington, D.C. 20551
david.o.lucca@frb.gov

Francesco Trebbi
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
ftrebbi@chicagobooth.edu

1 Introduction

After years of intentional opacity, central banks in advanced economies have exerted significant efforts to enhance the transparency of their communication to market participants (Mishkin [2007]). At least since the early 1990s, monetary policy communication has often been used to signal future policy rate decisions, for example by detailing the central banks' own economic projections on which these decisions are partly based. Forward policy communication can be thought of as an additional policy instrument that central banks can use to influence longer-term nominal interest rates beyond the conventional targeting of short-term interest rates through open market operations (Bernanke [2004], Woodford [2005]). Indeed longer-term yields reflect, albeit imperfectly, market expectations for the path of short-term rates, and thus they respond to credible communication of future policy decisions.[1]

With short-term interest rates close to the zero-nominal lower bound in many advanced economies for most of 2009, central bank communication has arguably become an even more important instrument. Several major central banks have actively used their communication to exert additional stimulus at a time in which conventional policy tools have been severely constrained. In the United States, the Federal Open Market Committee (FOMC) announced its intention to maintain "low levels of the federal funds rate for some time" in its December 18, 2008 policy statement. Perhaps even more forcefully so, other central banks, including the Bank of Canada, the Swedish Riksbank and the Reserve Bank of New Zealand, have also signalled future policy intentions through official communication. For example, after lowering its policy target to essentially zero, the Bank of Canada in its April 21, 2009 statement announced its conditional commitment to "hold current policy rate until the end of the second quarter of 2010."

Although the theoretical contributions to the analysis of information and communication have been significant in the economic literature, the empirical research has not produced comparable insights, partly because of the complexity involved in quantitatively evaluating flows of verbal information in a manner that is at the same time objective, intuitive, and replicable.[2] This paper attempts to fill such gap by advancing a class of automated measures of monetary policy communication, and it applies these measures to "FOMC statements", which are released by the Committee after its monetary policy meetings. These statements are of particular interest for two separate reasons. First, the texts represent an almost-ideal set of observations for an empirical analysis of communication: The structure of the text is fairly comparable over time; the statements are available for a relatively long period of time; finally, their release dates are fairly evenly spaced in time. Second, according to popular financial press and findings of previous literature, financial market participants pay close attention to their content, and changes in their wording elicit significant reactions in U.S. and other

[1] The imperfect response of long-term yields owes to the presence of term premia due to time-varying risk or other factors in the pricing of long-dated assets (see for example, Fama and Bliss [1987]).

[2] For a review of recent literature on central bank communication see Woodford [2005] and Blinder, Ehrmann, Fratzscher, Haan, and Jansen [2008].

financial markets.[3]

Text and words are not readily quantified in terms of intensity and direction of meaning, which in what follows will be referred to as semantic orientation. Interpretation of non-quantitative information is naturally subjective, and the same set of words can have very different meaning and intensity depending on the context of use and reader. In this paper, we borrow a set of tools from computer science and computational linguistics specifically designed to address these measurement issues and based on an intuitive, but information-theoretic, principle. We present two classes of automated scoring algorithms. The first score—the Google semantic orientation score—is directly based on the text of the FOMC statements, and is calculated using information from the Internet via Google-engine searches. The second score—the Factiva semantic orientation score—is constructed using discussions of FOMC statements from newspapers, journals and newswires that are included in the Dow Jones Factiva news database on days of announcements.

The construction of the Google semantic orientation score works as follows. Given two words representing opposing concepts (an antonymy, say, "hawkish" versus "dovish"), the semantic orientation of a sentence x (say, "Pressures on inflation have picked up") is measured by the relative frequency with which the string x and the word "hawkish" jointly occur, and the frequency with which string x and the word "dovish" jointly occur. If the string x co-occurs more often with the word "hawkish" than with the word "dovish", then it seems intuitive to attribute to that sentence a relatively more hawkish score (and vice versa). Contributions in the linguistic literature have provided an information-theoretic foundation to this approach.[4] Since it is not possible to directly compute joint frequencies of co-occurrence in the "population" of Internet webpages (at least unless massive computing resources are employed), we empirically implement the score using searches on Google. Indeed, hit counts on joint searches (for example, a search of the string x and the word "hawkish") represent empirical estimates of the population frequencies. Although the Google-based semantic score uses an extremely large set of information, it comes at the cost of not accessing potentially valuable information on the webpages, and on relying on the Google's proprietary algorithm to construct the hit counts of web-searches (not publicly disclosed). We therefore supplement the Google-based score, with another implementation measured on news from Factiva that we can directly access and analyze.

In the Factiva-based analysis we first subset news from the database that have headlines involving the Federal Reserve or the FOMC, around times of FOMC meetings and record all the sentences in the news. We then automatically analyze the text to construct a measure of semantic orientation of the FOMC announcement. Following the reasoning described above, given a sentence s from

[3] See for example Bernanke, Reinhart, and Sack [2004], or the recurring Wall Street Journal column " Parsing the Fed", which analyzes the content of each sentence of the statement relative to the most recent one "for clues about where interest rates may be headed".

[4] The concept of pointwise mutual information (PMI) employing information retrieval (IR) is discussed in the methodological part of the paper and, more in detail, in the companion Appendix to this paper. Relevant references in the linguistics literature include Church and Hanks [1990], Turney [2001, 2002] and Turney and Littman [2002].

a news article in the set of all text, the occurrence of the word "hawkish" as opposed to the word "dovish" within s justifies the attribution to s of a relatively more hawkish score.[5] By comparing the frequency of different antonyms—i.e. the mutual association of word pairs such as "hawkish/dovish", "loose/tight" to words such as "Fed or "interest rates"—we can generate scores for the large universe of news recorded in the Factiva database immediately before and after each FOMC announcement. Based on these we construct measures of semantic orientation and of (unanticipated) changes in these measures around the announcements.

After building these two automated scores, we study their properties as measures of monetary policy communication in a high- and low-frequency identification analysis. The sample starts in May 1999, when the FOMC begun systematically releasing statements after all of its policy meetings. We find that yields on short-term Treasury securities mainly respond to unexpected changes in the fed funds target rate during narrow time windows around the release of FOMC announcement. Instead, yields on longer-dated Treasuries only react to changes in the content of the statements, with 2-year Treasuries displaying the most pronounced yield responses.

In lower frequency data, we analyze the relation between the semantic scores and short-term rates within both a univariate and a vector autoregression (VAR) model.[6] The univariate model uses the semantic scores to directly forecast short-term rates—the federal funds and the 3-month Libor rate—at different horizons while conditioning on the information available to investors about future rates right ahead of the policy announcements.[7] The VAR model, instead, includes the federal funds rate and the semantic scores, as well as measures of inflation and economic activity. Parameter estimates of the univariate model imply that the scores have predictive power for short-term rates up to two years out. According to estimates of the VAR model, a one standard deviation increase in the scores implies a hump-shaped response of the federal funds rate with a peak of about 30 basis point after about one year. We then estimate the response of Treasury yields to the shocks identified in the VAR. Based on the forecast error variance decompositions, we find that the semantic scores have been more important than actual policy rate decisions as determinants of Treasury yields in the sample, consistent with the high-frequency results.

Finally, in order to analyze the type of information contained in the statements, we decompose realized policy rates with a forward-looking Taylor rule estimated using real-time forecasts of the output gap and inflation as proposed by Orphanides [2001]. Based on this decomposition we find that the semantic scores contain significant information regarding both the predicted and the

[5]As we describe in more detail in Section 2, the Factiva approach allows for a high degree of precision in the textual searches. For example, we can easily accommodate for negations of our matches in the analysis (for example a match of "not hawkish").

[6]In the paper we find that the Google-based scores is somewhat noisier that its Factiva counterpart. The low-frequency results discussed here refer to this latter score. See the conclusions for more discussion on this point.

[7]As implied by quotes on futures contracts delivering at the corresponding time horizon. We consider Libor rates to study the scores' predictive power for future short-term rates beyond the first few months. Indeed, while the liquidity of Eurodollar futures contracts—settled on future Libor rates—that expire beyond one-year is relatively high, the liquidity of federal funds rate futures drops sharply for expirations beyond the first few months.

residual component of the rule-implied interest rate decisions. In addition, the semantic scores lead the residual component of the Taylor-rule by a few quarters.

These findings based on our automated semantic measures support the view that the FOMC alters the content of the statement several months ahead of taking policy rate actions, and consistently longer-term nominal Treasury yields respond to changes in the content of the statements. The automated approach to central bank communication that we present in this paper is new in the economic literature and we think that it has several advantages relative to previous literature. First, it does not rely on subjective ratings of text by researchers, like for instance in Romer and Romer [2004] or Bernanke, Reinhart, and Sack [2004]. At the same time, by specifying an ex-ante metric along which analyzing the content of the statement—in particular, one can focus on the degree of "hawkishness" of the statement as predictor of future policy rate hikes—we depart from black-box methodologies, such as latent analysis methods, which deliver findings that are hard to interpret economically, and are often silent about policy communication prescription.[8]

The remaining of paper is organized as follows. In Section 2 we present the methodological description of the automated measures set forth in this paper, and apply them to the FOMC statements (with additional detail presented in the companion Appendix). In Section 3 we present the data used in the empirical analysis. In Section 4 we investigate the effects of communication on asset yields using a high-frequency identification. In Section 5 we analyze the low-frequency properties of our linguistic scores: we first evaluate their forecasting power in a univariate setup; we then analyze the empirical link between the linguistic scores and the systematic and non-systematic components of future and current policy decisions, as determined by Taylor rules; finally, we analyze the relation between the scores, policy actions and Treasury rates in a recursive VAR model. Section 6 concludes.

2 Automated measures of the FOMC statement

Over the course of the past decade, FOMC statements have arguably been among the most important means used by Committee members to communicate to investors about monetary policy. The statements are short in length: In our sample, the core of the statements is composed on average of about six sentences, each of which about 25 words long, expressing succinctly the FOMC's rationale for the most recent policy action, or lack thereof, and an assessment of the risks to its goals of "price stability and maximum sustainable employment"going forward.[9] In order to put the the measurement approach in a sharper perspective, we briefly discuss the role of FOMC statements,

[8]See for example Boukus and Rosenberg [2006] for an application of latent semantic analysis to FOMC minutes. Gürkaynak, Sack, and Swanson [2005] apply latent factor analysis to yield responses around FOMC announcement to indirectly measure the impact of communication on yields.

[9]The sample period starts in May 18, 1999 and ends in December 16, 2008. See Section 3 for additional detail. In the paper we will refer to the FOMC statement as its "core" text, that is, excluding the preamble describing the policy rate action taken at the meeting, and the concluding list of the voting members' roll call.

and of central bank communication more in general, as monetary policy instruments. We then turn to the measurement approach.

Central bank communication Central banks in developed economies often use their official communication to influence market participants' views about the likely path of future policy actions, and to align these views to their own. Communication can be used to improve the understanding of long-run policy objectives, helping central banks to achieve their long-term goals, most notably anchoring inflation expectations (Bernanke [2004]).[10] As agents' expectations concerning future policy rate moves help determine the paths of aggregate prices and quantities, central banks can achieve Pareto-superior equilibria when they control these expectations by committing to specific policy paths (Woodford [2005]).[11]

In practice, central bank communication increases the effectiveness of monetary policy by influencing long-term interest rates beyond the immediate setting of short-term policy rates, as long-term rates depend not just on the current, but on the entire expected path of short-term rates up to term premia. Because long-term rates generally have a very important role in households' and businesses' economic decisions—for example, through mortgage or corporate bond rates—by influencing market expectations central bank communication can enhance the effectiveness of their policies.[12] When policy rates are constrained by the zero-nominal bound communication becomes an even more important tool. For example, as shown by Eggertsson and Woodford [2003], in standard macroeconomic models with a complete set of frictionless financial markets, communication about the future policy path (so long as it is perceived as being credible) is among the very few policy instruments available to central banks to exert monetary stimulus at the zero-nominal bound.[13]

Measurement of the FOMC statement Consistent with an interpretation of central bank communication as an expectation management tool, we construct a set of measures of the content of FOMC statements that attempt to extract information about future policy rate actions. There

[10] Among other important reasons discussed in Bernanke [2004] for central bank communication, and transparency more in general, are that systematic communication increases the accountability of monetary authorities, a particularly important goal given their central role and political independence. In addition, because unexpected policy actions come with large asset price reactions, as well as with large "winners" and "losers", communication may help improve the overall stability of the financial system.

[11] With time-varying objectives and preferences that are not directly observed by other agents, central banks can use communication to signal the future policy path, therefore aligning agents' expectations to the central bank's preferred policy path. Of course, under rational expectations, future policy moves have to be confirmed in equilibrium, implying an empirical correlation between words and future actions. We find significant evidence of such correlation in Section 5.

[12] Short-term rates clearly matter directly for economic decisions. For example, the average prime rate on business loans in the U.S. is priced off the intended federal funds rate with a spread of 300 basis points (H.15 Federal Reserve Statistical Release). The term of business loans is in general shorter than the term on other loan categories, most notably residential mortgage and commercial real estate loans.

[13] In particular, Eggertsson and Woodford [2003] show that nonconventional policies, such as purchases of financial assets by central banks in the open market, are neutral in many standard macroeconomic models. In these models, long term yields are only pinned down by the expected path of short term rates and monetary policy cannot affect term- or risk-premia by other means. Bernanke, Reinhart, and Sack [2004] present some empirical evidence against these stylized predictions.

are at least two more reasons to focus on this dimension of communication in addition to what just discussed. First, it is directly suggested by the "balance of risk" of the statements.[14] Second, as evidenced by popular financial press, investors attempt to extract information about future policy actions from the statements, therefore guaranteeing sufficient data on news outlets and the Internet that we can draw upon to construct our automated measures. In what follows, we define measures of the policy stance, or policy "hawkishness", and its intensity based on the FOMC statements. The measures should produce a high score for a *hawkish* statement—one that hints to a possible increase in the target funds rate—and a low score for a *dovish* statement, that is, one that implies a decrease in the target rate.

The inherent difficulty of measuring words' meaning, discourse orientation, and intensity is the primary challenge in constructing these measures. For the sake of concreteness, suppose that we were set to analyze the information content regarding future rate decisions implied by the two phrases: "Pressures on inflation have picked up", statement of March 22, 2005—call this string of text X; and "Inflation pressures seem likely to moderate over time", December 12, 2006—call it string X'.[15] Although it seems natural to interpret the former phrase as being more hawkish than the latter, no clear metric exists prima facie to assess the two.

A heuristic measurement approach In order to emphasize the advantages of our automated scores, we discuss next a simple approach to go about the problem. Let us assign to each sentence a subjective score. For instance, consider the following scheme, which we will call heuristic index, or score:

$$HI(x) = \begin{cases} 1 & \text{if the sentence indicates, or suggests, an increase in inflation;} \\ -1 & \text{if the sentence indicates, or suggests, a decline in economic activity;} \\ 0 & \text{if neutral.} \end{cases} \tag{1}$$

According to the operator defined in (1), the score $HI(X)$ would clearly be a 1, whereas $HI(X')$ would possibly be a 0. A heuristic approach such as the one just described has advantages and

[14] The FOMC statement has included both direct and indirect references to future policy rate actions. Until January 2000 the statement contained an explicit reference to subsequent policy moves called policy "bias". This was later replaced with a "balance of risks" that only indirectly discusses policy moves through an assessment of the weights given to the objectives of price stability and growth. A direct reference to policy rate actions was reintroduced in statements during the "zero-nominal bound" periods of 2003-2004 and after December 2008, for example, by noting "that policy accommodation can be maintained for a considerable period" in the August and December 2003 statements. Kohn [2005] discusses potential advantages of providing only indirect references to future policy actions. According to his view, an indirect discussion of future policies through policy objectives provides a clearer indication that the commitments to future policy moves are state-contingent, rather than unconditional binding promises that cannot take into account future evolution of policy-relevant variables.

[15] The strings X and X' are, respectively, part of the following two sentences: "Though longer-term inflation expectations remain well contained, pressures on inflation have picked up in recent months and pricing power is more evident."and "However, inflation pressures seem likely to moderate over time, reflecting reduced impetus from energy prices, contained inflation expectations, and the cumulative effects of monetary policy actions and other factors restraining aggregate demand."

shortcomings.[16] It is an intuitive and a simple measure of the orientation of a sentence. However, it coarsely approximates intensity and relies on an arbitrary and subjective judgment by the researcher, and it is, thus, difficult to interpret or replicate across scorers. Consider, for instance, the alternative interpretation of the score for X' as being equal to -1 on the grounds that lower inflationary pressures can be associated with a decline in aggregate demand, and that this scenario might be unwarranted when inflation is below a central bank's target. In addition, the categories in (1) might not always be mutually exclusive as, for example, in the case of a discussion of stagflation.

Putting these concerns aside, one might ask how accurately such a score describes the monetary policy stance of FOMC statements. We discuss this issue in more detail in the next sections, but as a preview, we construct a time series of the heuristic score, by applying the scheme (1) to the set of sentences in each FOMC statement in our sample, and then by averaging within each statement.[17] Figure 1 shows the heuristic score alongside the intended (target) federal funds rate, and the rate implied by the fourth quarterly Eurodollar futures contract (taken 20 minutes after the FOMC announcement), which, as discussed in Section 3, is a market based measure of short-term interest rates at the one-year horizon. From a descriptive standpoint, the heuristic score appears to lead the target rate by a few quarters, consistent with the idea that the score measures information about future policy actions. Statements with a dovish stance appear to lead subsequent policy rate cuts, while statements with a hawkish stance have been followed by subsequent rate hikes. The heuristic score is also fairly correlated with the contemporaneous level of the Eurodollar futures rate (about 50 percent in levels), although the score has been somewhat more volatile than this rate at times.

The automated measurement approach The novel application that we propose in this paper is to assign to each sentence an objective and automated score able to capture the semantic orientation of the statement, or one of its parts, along a "hawkish–dovish" metric (or alternative metrics deemed appropriate—we experiment with six in total). Although a relatively new problem in economics, the unsupervised and automatic measurement of the intensity, or semantic orientation, of text is commonplace and a long-standing issue in computational linguistics and statistical natural language processing—scientific fields at the intersection between computer science and linguistics (see the book by Manning and Schütze [1999] for a review).[18] As such, there is a considerable variety of alternative approaches aiming at quantitatively "evaluating" the meaning of language.

[16] See Romer and Romer [2004] and Bernanke, Reinhart, and Sack [2004] for applications of subjective measures in the analysis of the Fed's policy stance and communication using FOMC minutes/transcripts and statements, respectively.

[17] The heuristic score reported is the consensus on the analysis of each statements by the three reviewers (including the authors). We limit the number of scorers due to the inherent subjective nature of the score. This notwithstanding, the relatively high concordance in assessing the orientation of statements across the different scorers reveals that for several phrases there seems to relatively little ambiguity of interpretation of rule (1).

[18] There are already few interesting applications of computational linguistics in the economic literature, however. For an application of latent semantic analysis to FOMC minutes see Boukus and Rosenberg [2006]. Other examples include Stock and Trebbi [2003], Antweiler and Frank [2004], Tetlock [2007], and Gentzkow and Shapiro [2006] and Demers and Vega [2008]. Intuitive linguistics indices, such as word counts, have been occasionally the focus of research in monetary economics, for example in Gorodnichenko and Shapiro [2007].

We follow two different methodologies to validate the analysis from different perspectives. A first approach—the Google semantic orientation score—relies on estimating the systematic co-occurrence of concepts retrieved from webpages, an idea initially proposed by Church and Hanks [1990], and first applied to the analysis of semantic orientation through information retrieval on the Internet by Turney [2002]. As described below, the directness of the approach makes it relatively easy to implement and transparent to interpret as compared to other methodologies proposed in the computational linguistics literature.[19] Our second approach—the Factiva semantic orientation score—relies on analogous principles of information retrieval, but it is based on the discussion of FOMC statements in news outlets. This approach has the main advantage of letting us directly access the corpus of text used to calculate the semantic score (the whole universe of documents in the Dow Jones Factiva database released right before and after a given FOMC announcement), drastically increasing the precision in measurement.

We will start from a description of the first measure. Due to space constraints, we present here a general characterization of the scores and relegate the bulk of the implementation details to the companion Appendix.[20]

2.1 The Google semantic orientation score

Assume that the metric that we wish to define can be properly characterized along a simple antonymy, say for instance scoring a string of text x—either a word, a phrase, or a sentence—on an "hawkish-dovish" scale. That is, we wish to create a score defined over the real line whose values depend on the hawkishness of the given string of text x as implied by a reference corpus of text, in this case, the universe of webpages. We begin by defining a measure of association between concepts. If the meaning of x can be commonly interpreted as hawkish, then x and the word "hawkish" should show a degree of positive statistical dependence in a sufficiently large corpus of text. In other words, the string x and the word "hawkish" should appear in a corpus of the text with a joint frequency, $\Pr(x \,\&\, \text{hawkish})$, which is greater than if the two strings were statistically independent concepts, in which case the joint frequency would be equal to the product of the marginals, or $\Pr(x)\Pr(\text{hawkish})$. The Pointwise Mutual Information (PMI, Church and Hanks [1990]) between the string of text x and the word "hawkish" is defined as:

$$PMI(x, \text{hawkish}) = \log\left(\frac{\Pr(x \,\&\, \text{hawkish})}{\Pr(x)\Pr(\text{hawkish})}\right). \tag{2}$$

PMI is a central concept in information theory, and it is derived from the joint entropy of two random variables (see Manning and Schütze [1999] for a review). Specifically, given two elements, PMI is a log-ratio indicating the amount of information that it is possible to gather about one

[19] For instance methods that require learning algorithms, such as the one employed by Hatzivassiloglou and McKeown [1997] in the study or semantic orientation of adjectives, or methods involving factor decompositions that are difficult to interpret, such as latent semantic analysis (Landauer and Dumais [1997], also see Boukus and Rosenberg [2006] for an application to monetary policy). See Turney [2001] and Turney and Littman [2002] for a comparison across the different approaches.

[20] Available at: `faculty.chicagobooth.edu/francesco.trebbi/research.html`

element of the message when the other message is observed.[21] A measure of the relative degree of association between the string x and the word "dovish" can be computed accordingly as in (2), hence obtaining the degree of "dovishness" that we can attribute to x. In order to obtain a measure of orientation we consider the relative PMI measure between the two polar concepts, and obtain a theoretical score of semantic orientation (SO) of string x, based on (2), as:[22]

$$SO(x) = PMI(x, \text{hawkish}) - PMI(x, \text{dovish}).$$

The Internet represents a very large corpus of text from which it is possible to obtain empirical frequencies of each string of text in a statement, and the words "hawkish" and "dovish". Since it is unfeasible to directly compute the joint frequencies of co-occurrence in the actual population of webpages (at least without devoting huge computing resources), we empirically implement the information retrieval (IR) process through hits counts on the search engine Google.[23] The feasible estimator of the semantic orientation score obtained by information retrieval on Google using the "hawkish-dovish" word pair is:

$$\begin{aligned}\overline{GSO}^h(x) &= \log\left(\frac{\text{hits}(x \,\&\, \text{hawkish}) * \text{hits}(\text{dovish})}{\text{hits}(x \,\&\, \text{dovish}) * \text{hits}(\text{hawkish})}\right) \\ &= \log\left(\frac{\text{hits}(x \,\&\, \text{hawkish})}{\text{hits}(x \,\&\, \text{dovish})}\right) + \xi, \end{aligned} \tag{3}$$

where $\text{hits}(q)$ is a function assigning the number of hits in the search of query q and ξ is a constant that is independent of the specific string x being searched.[24] The superscript "h" in $\overline{GSO}^h$ denotes the antonymy "hawkish-dovish" used in building the score. The $\overline{GSO}^h$ score is defined over $(-\infty, \infty)$ and is increasing in the degree of hawkishness of the string of text x. Computing the scores associated with the strings X and X', which we presented earlier in this Section, is straightforward by implementing six searches in Google.[25] For the hawkish sentence X we obtain

[21] As such, both computer scientists and linguists employ PMI as a measure of association between words, word pairs and strings of text. In computing (2) we employ the base e instead of the base 2 as customary in the literature (Turney [2002]). The base is immaterial, as the two measures are equivalent up to a constant.

[22] See Turney [2002] and Turney and Littman [2002].

[23] Available at: `www.google.com`. We make use of the University Research Program for Google Search for the necessary high-volume sequential access. This search engine coverage can be thought as being virtually complete, and its index of webpages is the largest available—it included 8,168,684,336 webpages in September 26, 2005 according to the New York Times. Turney [2002] implements his searches on `www.altavista.com`, another popular search engine because of the availability (at the time) of a NEAR operator to condition joint occurrences to be in a ten words radius on searched webpages. This operator is no longer available on Altavista and it is also not available on Google. Each search individually run on Google is rerouted to a specific data center depending on Internet traffic. Since each data centers caches are slightly different, we constrain our searches within the same data center by conditioning the search on a common data center's IP address. We experimented with several centers obtaining similar results. We also rerun all our searches leaving the IP address unspecified. Although mildly more noisy, the results were also unaffected. Finally we also ran searches on Altavista, which currently implements searches through the Yahoo! search engine, obtaining similar results.

[24] Searches for units x containing more than one word are run as "phrase searches", that is, they are enclosed in quotation marks (see `www.google.com/support/websearch/bin/answer.py?answer=136861`).

[25] In the Appendix we detail the type of restrictions that should be applied in order to direct searches on identical

the higher score $\overline{GSO}^h(X) = .98$, and the lower score: $\overline{GSO}^h(X') = -.53$ for the relatively more dovish sentence X'. This example is representative of how an unsupervised, automated algorithm such as the SO-PMI can approximate a subjective interpretation of a string of text along the hawkish–dovish dimension.[26] Before further discussing the measurement of the FOMC statement using the Google-based score, it is important to note that the antonymy "hawkish-dovish" (or others that we consider below) is arguably used in contexts that are different from discussing monetary policy (such as documents related to international relations). Because the relative frequencies with which the two words occur in the discussions of monetary policy could differ from those in general language use, the constant ξ is probably measured with error, adding a constant "level effect" to the measure. We directly abstract from this level by considering the measure:

$$GSO^h(x) = \overline{GSO}^h(x) - \xi, \tag{4}$$

in what follows, therefore avoiding to interpret the level of the score. This does not restrict our empirical analysis, however, which relies on using the score in first differences that do not depend on the constant terms.

We implement the scheme (4) on each sentence of an FOMC statement as follows. We first apply a Brill [1994] part-of-speech tagger to the text, which is a natural language processing algorithm used to automatically classify and tag words in the lexical categories of nouns, adjectives, verbs, adverbs, pronouns and coordinating conjunctions. After this pre-processing, we apply an automated routine to obtain groups—chunks—of related words based on the lexical tags corresponding to either verbal, noun or adjectival phrases. We finally run searches on sub-sentences obtained by joining five (or less) of these chunks.[27] To obtain the search units x, the algorithm moves on a rolling window for sentences including more than five chunks, and uses the whole sentence otherwise. After obtaining the scores for each search unit x, we average $GSO^h(x)$ over all x in the statement to obtain a score for statement t.[28]

Because we cannot directly measure expectations for the content of the statement on the Internet, in

Google data centers and side-step automatic Internet traffic optimizers.

[26] The Internet evolves continuously over time. Pages are substituted and dropped from Google caches over time and its index algorithm is run every day. This implies that searches executed at different moments, even on the same data center, may differ. We run our searches in April 2007, August 2007, and May 2009 and found a correlation across hits of above 80 percent, indicating a substantial, although not perfect, degree of persistence. One general interpretation is that the strength of association between concepts is persistent and does not change as fast as the Internet over time.

[27] We include only searches longer than three words, to exclude incidentals and sentences for which not clear semantic orientation can be defined. For instance running a query for the sub-sentence "in any event," or "the committee believes" would not be meaningful in our context. We experimented with three, four, five, and six chunks and obtained the best fit employing the four and five chunks. We report the analysis performed joining five chunks, although the results in the following sections are robust to alternative sub-sentence lengths. Employing searches directly on whole sentences, however, did report zero hits very frequently (hence a problem for the logs) and resulted in large number missing observations and noisy measurement. The robustness results are available from the authors upon request. The automated searches and natural language processing of the text is implemented in Python 2.5, using routines from Liu [2004] and Bird and Loper [2006]. All details are in Appendix.

[28] With a slight abuse of notation we also indicate with the release date t the statement released on that date.

the empirical analysis below, we approximate the unexpected change in the content of the statement at date t as the difference between the semantic orientation score at meeting t and the score at meeting $t-1$:

$$\Delta GSO_t^h = GSO_t^h - GSO_{t-1}^h. \tag{5}$$

Of course, such an approximation will be true only when the score evolves according to a random walk, while, otherwise, the approximation will imply a measurement error in the construction of the shocks to the content of the statement.[29] We return to these issues below in the paper.

For simplicity of exposition we have focused thus far on the antonymy "hawkish-dovish" to construct the Google-based score. These two words are particularly pertinent for our analysis as they are frequently used in the financial press and by other market commentators when discussing U.S. monetary policy. More generally, the SO-PMI measure can be extended to consider multiple sets of antonyms simultaneously. As an alternative to the GSO^h score, we consider a Google-based score implementation using six pairs of words in total. In this implementation, words associated with positive rate changes form the list $\mathbf{P} = \{$hawkish, tighten, hike, raise, increase, boost$\}$, while words associated with negative changes in policy rates movement form the list $\mathbf{N} = \{$dovish, ease, cut, lower, decrease, loose$\}$.[30] We define the score GSO^e, as the semantic orientation score based on the extended set of **P** and **N** antonymies. The GSO^e score adds the total number of joint hits of x and each of the words predicting positive (respectively, negative) interest rates movement and uses this sum in the numerator (respectively, denominator) of the argument of the logarithm in (4). This definition can be justified in thinking of the words within each of the sets **P** and **N**, as having the same meaning, that is, as being synonyms along our metric of semantic orientation.

By focusing on a set of synonyms, the GSO^e score captures additional information as compared to GSO^h, thus increasing the coverage of the score, that is, the number of search hits per sentence, and the fraction of sentences for which we can find search hits in Google. However, by leveraging on word pairs with as much pertinence to current policy actions as to future policy moves, more noise is added to the measure.[31] In particular, measurement error is a problem that compounds when dealing with differences as in (5). Nonetheless, in levels the fully automated score GSO^e appears reasonable as a measure of communication, and leads the policy rate by about two quarters, as reported in Figure 2, as accurately as, if not more than, the heuristic index. The correlation of the GSO^e score and the fourth Eurodollar futures implied rate is about 40 percent (Table 1).

[29] The first order serial autocorrelation of the GSO^h score is about .7 in our sample.

[30] Google searches automatically include hits based on related matches such as "tight", "increasing", etc.

[31] The words hawkish and dovish are particularly pertinent to the analysis of central bank stance and forward policy, since they are often used in the context of the FOMC's near-future interest rates behavior. Words like hike or cut instead often appear in the discussion of both current hikes/cuts of target rates and future hikes/cuts, making more ambiguous their pertinence as a measure of forward policy.

2.2 The Factiva semantic orientation score

The Google semantic orientation score relies on the Internet as the corpus of text on which the joint frequencies that form the score are estimated. Because we can only access the text of the corpus of webpages indirectly through Google searches, we have a rather limited control over which specific texts the search is run over, the specific time periods of reference, or the relevance of the matches obtained from the search engine. In addition, Google does not publicly disclose the algorithm it uses to calculate and approximate the count of hits of a given search. These restrictions can be avoided by implementing a more standard IR protocol on a corpus of text for which we can directly access the text underlying the corpus. This is the main advantage of the Factiva-based score that we now present.

To construct this score we use discussions of FOMC announcements from newspaper, magazine, newswires and newsletters that are included in the Dow Jones Factiva database, a leading provider of business and financial news.[32] The original documents range from very short pieces of newswire information to long newspaper articles and commentaries. To implement the score we search all sources available worldwide in English, for articles with headlines involving the words "Federal Reserve", "Fed" of "FOMC", around times of FOMC meetings and record all the sentences in the database that match this criterion. We select all these articles on a 3-days window around the FOMC meeting starting on the day before, and ending on the day after, the announcement. This lets us focus on information as pertinent as possible to each given policy announcement. All documents are then subdivided into sentences and assigned a unique identifier. This set of sentences constitutes the corpus of text, $\mathbf{T}$, on which we run our text searches to form the Factiva-based score.[33] The total number of sentences in $\mathbf{T}$ is $1,302,977$ or about $15,512$ sentences per statement. Let us indicate with $\mathbf{T}_t$ the set of sentences in news articles around release date t and hence pertaining to statement t.

Although this corpus of text is relatively large, it is by many orders of magnitude smaller than the Google's search cache. Thus, rather than trying to find direct matches of strings of text from each statement, we directly measure the semantic orientation of sentences referring to the statement or interest rate decisions as a proxy for the stance of policy. Let us define as relevant for the FOMC meeting discussion a sentence s if it contains any word in the list $\mathbf{R} = \{$Rates, Policy, Policies, Statement, Announcement, Fed, FOMC, Federal Reserve$\}$.[34] Define as $I[s, \mathbf{W}, \mathbf{W}']$ an indicator function that takes value 1 if sentence s contains at least a word from list $\mathbf{W}$ and at least a word from list $\mathbf{W}'$, and 0 otherwise. [35] We compute the Factiva semantic orientation score

[32] Factiva includes a collection of more than $25,000$ news sources including The Wall Street Journal, the Financial Times, Dow Jones Reuters, and the Associated Press.

[33] We implement the search using "regular expressions" in Python.

[34] About 40 percent of all sentences in the corpus are relevant according to this criterion.

[35] An example of $I[s, \mathbf{R}, \mathbf{P}] = 1$ match is the sentence "*FED WATCH: Economists Expect At Least One More Rate Hike*", the title of a Dow Jones Capital Markets wire on January 31, 2006.

for statement t as:

$$FSO_t = \log\left(\frac{\sum\limits_{s\in\mathbf{T}_t} I\left[s,\mathbf{R},\mathbf{P}\right]}{\sum\limits_{s\in\mathbf{T}_t} I[s,\mathbf{R},\mathbf{N}]}\right) \tag{6}$$

where $\mathbf{P}$ indicates the set of words associated with positive target rates movement (respectively, $\mathbf{N}$ for negative). Equation (6) is the log-ratio of the frequencies of relevant sentences suggesting policy rate increases as opposed to relevant sentences suggesting rate decreases. The role of (6) mirrors the type of co-occurrence that (4) estimates through Google hit counts. We calculate the score FSO^e using the list of antonymies that defined the score GSO^e when also including as matches in (6) strictly equivalent words in the sets $\mathbf{P}$ and $\mathbf{N}$.[36] In a similar fashion to the Google-based GSO^h score, we also consider a Factiva "hawkish-dovish" index, FSO^h, by collapsing the $\mathbf{P}$ list to the word "hawkish" only, and to the word "dovish" for $\mathbf{N}$.

Let us further indicate with $\mathbf{T}_{t^-} \subset \mathbf{T}_t$ the set of sentences in the Factiva corpus in the 3-days window that precede the FOMC announcement and with $\mathbf{T}_{t^+} \subset \mathbf{T}_t$ the set of sentences in news articles released after the announcement. We compute unexpected changes in the stance of the FOMC announcement at t as the difference between the Factiva semantic orientation score based on news released before and after the announcement:

$$\Delta FSO_t = FSO_{t^+} - FSO_{t^-}. \tag{7}$$

Note that the calculation of the difference (7) reflects the high degree of precision obtained by directly accessing the underlying text to determine the exact time of the information included in constructing (6). Such precision is unfeasible for the Google-based score (at least under the current state of the search technology), but easy to achieve in Factiva. For the expanded set of antonymies, it is important to notice here that both (6) and (7) are likely to capture discussion concerning not just policy communication, but the policy action taken at the meeting. Fortunately, having full control over the text search, alleviates in part this problem (in addition to controlling for the immediate policy action in the empirical specifications). In particular, we compute FSO_{t^+} by removing all instances of matches in the past tense for verbs, thus avoiding discussions of the most recent or past policy action (at t) and thus focusing on future policy moves (and similarly for FSO_{t^-} although only discussions of past, rather than the immediately forthcoming action at t, can be excluded this way). We finally refine our measures by excluding from the set of joint matches direct negations of the words included in the list of antonymies (for example, "not hawkish") and include direct negations of the opposite (for example, "not dovish" for "hawkish"). Notice that such a degree of precision cannot be achieved within Google, where conditions of relevance or vicinity of negations cannot be reliably imposed within the search protocol.

As reported in Figure 3, the Factiva automated score FSO^e constructed with this algorithm leads

[36] For example "rates" is considered an equivalent to the word "rate". We achieve this by matching the roots of the words in each set, rather the exact words, using regular expressions in the search code.

the policy rate by more than two quarters, with movements in levels that track the rate implied by the fourth Eurodollar futures contract fairly accurately (a correlation of about 40 percent). The Google- and Factiva based scores defined on the extended set of antonymies also display a correlation of over 80 percent (Table 1). We describe the empirical properties of the Factiva- and Google-based scores in more detail in the rest of the paper, after discussing next the data used in the analysis.

3 Data

This Section describes the data used in the empirical analysis. The sample includes 82 FOMC statements starting in May 1999 and ending in December 2008.[37] The starting point marks the date in which the FOMC begun releasing statements after all meetings, irrespective of whether a change in the target federal funds rate was announced at that meeting or not. For each full calendar year in our sample, the FOMC released 8 statements following so-called "scheduled" policy meetings. Although the dates of scheduled meetings can shift by up to a couple of weeks from year to year, the precise schedule of these meetings is set (and communicated to the public) with about 6 months lead. Scheduled meetings are not equally spaced on calendar years, but have occurred almost every 1-1/2 months in our sample. In 2001, 2007 and 2008, the FOMC also released statements following unscheduled meetings. These statements were associated with either intermeeting policy rate changes–always cuts in our sample—or, after the onset of the financial turmoil in August 2007, with other policy actions aimed at relieving pressures in financial markets.[38] Because this paper focuses on communication about monetary policy rates, we only include statements for unscheduled policy meetings that discuss current or future policy rate decisions. These statements have always coincided with meetings at which changes in the current target federal funds or discount rate were announced. In addition, due to missing financial quotes after the September 11 terrorist attacks, the September 17, 2001 statement is not included in our regressions.

Interest rate response regressions In the next Section, we study interest rate responses on narrow temporal windows around FOMC announcements to interest rate decisions and communication. In this analysis we use high-frequency intra-day data to better isolate the impact of policy actions from other same-day events, such as economic data releases. We consider interest rate responses on time windows that are 30 minutes long, starting 10 minutes before and ending 20 minutes after the announcements. The dependent variables are basis point yield changes of on-the-run (benchmark) Treasury securities, including bills with maturities of 3- and 6-months, as well as 2-, 5-, 10-, and 30-years coupon-bearing Treasuries. In addition, we also consider implied rates on short and medium-dated Eurodollar futures contracts. These contracts are cash settled on realized

[37] The statements are available on the website of the Federal Reserve Board: `www.federalreserve.gov/monetarypolicy/fomc.htm`.

[38] For example on August 10, 2007, the FOMC statement noted that the Federal Reserve was providing liquidity to facilitate the functioning of short-term interbank funding market. Several of the new lending facilities aimed at addressing pressures in short-term funding markets were announced in 2008 using FOMC statements.

Libor rates at settlement dates.[39] Our explanatory variables include changes in policy rates and announcements. Following Kuttner [2001], the regressions include the *unexpected* component of policy rate decisions—hereafter, referred to as the monetary policy surprise—as measured by the change in the current-month federal funds futures contract, rescaled to account for the date of the meeting within each month.[40] As discussed in Section 2, for Google-based scores we construct a measure of unexpected shock to the content of the policy statements as the change in the current semantic score relative to the score of previous FOMC meeting. For the scores constructed on Factiva data, instead, the shocks are measured as differences between the content of news articles released approximately in the 36 hours after relative to those released in the 36 hours prior each FOMC announcements.[41]

Table 2 reports summary measures for the dependent and independent variables included in the regressions. As seen in the upper-panel of the Table, the standard deviation of Treasury yield changes range between about 5 and 6-1/2 basis points depending on the maturity of the Treasury security; the 2-year security is the security that displays the largest volatility. Changes in Eurodollar futures implied rates are slightly more volatile than Treasury yields (middle panel of Table 2), and are highly correlated with changes in Treasury yields of comparable maturities (not reported in the Tables). With a few exceptions, interest changes around announcements are fairly small in magnitude, pointing to the fact that yields already incorporate much of the actions and communication ahead of the actual announcements. The largest yield declines shown in the Table occurred following the 50 basis point intermeeting cut on April 18, 2001. The largest increase, instead, occurred following the January 28, 2004 meeting, reportedly due to the omission of the phrase "policy accommodation can be maintained for a considerable period" in the FOMC statement, which, according to market reports, was interpreted by investors as a signal that a policy tightening cycle could begin in the near future. The lower-panel of Table 2 and Table 3 present summary measures and pairwise correlations for the monetary policy surprise and semantic scores based on the hawkish/dovish and the extended set of antonymies discussed in the previous Section on information from Google and from Factiva (respectively, GSO^h, GSO^e, FSO^h and FSO^e). As

[39] The source of intra-day data is the internal database of the Federal Reserve Board. On-the-run Treasury securities are, for each maturity, the ones being most recently auctioned by the U.S. Treasury. These securities are more actively traded in the secondary market than their off-the-run counterparts. Eurodollar futures contracts are obligations for the seller to deliver fixed amounts of Eurodollar 3-months deposits at expiration (the contract is quoted as p=100-r, where r = 3-month Libor; all results in the paper refer to the implicit yield, r, rather than to the price,p). At each moment in time, price quotes are available for quarterly contracts expiring in mid-March, June, September and December for the following ten years (for each month, the delivery date is the second London bank business day before the third Wednesday of the month). For example, in August 2007, the second contract is December 2007, while the eighth contract is the June 2009. The liquidity of contracts expiring far in the future is fairly limited and so we only include the first eight contracts in our analysis.

[40] Federal funds futures contracts are priced on the the average effective federal funds rate for the month of expiration. The monetary policy surprise is calculated as: $(FF_{t^+} - FF_{t^-}) \cdot dm/(dm - d)$, where FF_{t^+} and FF_{t^-} are the futures federal funds rate after and before the FOMC announcement, respectively. The scaling factor, $dm/(dm - d)$, adjusts for the averaging effect of the federal funds futures rates (dm denotes the total number of days in the month, and d is the day of the month in which the meeting takes place). See Kuttner [2001] for more details.

[41] For unscheduled meetings, the difference is Factiva scores refers to the prior meeting score as it is the case for Google scores in general.

shown in Table 2, the FSO^h score is only defined for about two-thirds of the observations included in the sample. With the exception of the GSO^e, changes in all semantic scores are uncorrelated with monetary policy surprises, highlighting how these measures are likely capturing information, which is unrelated to current policy actions. As discussed in Section 2, the GSO^e score is likely measured with error as its level is likely influenced and partially reflects contemporaneous rate decisions taken at the meeting. This is due to the difficulty in matching with precision the extended set of antonymies through Google searches.

For an easier interpretation of the coefficients in Section 4 we standardize changes in all semantic scores—imposing a zero mean and unit standard deviation—so that the units of the coefficients are expressed as basis point changes of each dependent variable per unit standard deviation increase in the scores.

Univariate interest rate forecasting regressions These forecasting regressions assess the predictive content of changes in the semantic scores around FOMC announcements for future realized short-term interest rates. The regressions include as controls, the monetary policy surprise as defined above, the slope of the Treasury yield curve measured as the difference between yields of 10- and 2-year Treasuries using quotes 10 minutes ahead of the FOMC announcements, a credit spread between 10-year BBB-rated corporate bonds and Treasury yields at the close of the previous business day, a dummy for NBER dated recessions.[42] The regressions also include futures rates as of 10 minutes before the announcement implied by the first 8 quarterly Eurodollar and the first 6 federal funds futures. The dependent variables in the regressions are chosen to match the settlement rates of the futures contracts included in the regressions: Monthly averages of the federal funds rate for the "federal funds rate" regressions and the Libor rate on settlement dates for the "Eurodollar futures" regressions.

Taylor rule analysis In this analysis we mainly follow the forward looking Taylor rule implementation of Orphanides [2001, 2003] and use mid-quarter Greenbook forecasts of the output gap and output deflator, prepared by the staff of the Federal Reserve Board on the week preceding each scheduled policy meeting. These data are released to the public with a 5-year lag, and are available on the website of the Federal Reserve Bank of Philadelphia. We supplement the last 5-year of missing data with real-time measures of output potential from the Congressional Budget Office and forecasts of GDP and of the GDP deflator from the Survey of Professional Forecasters (SPF). The SPF data are also available on the website of the Federal Reserve Bank of Philadelphia. The federal funds rate and semantic scores included in the analysis are average quarterly levels.

VAR analysis The VAR is estimated at monthly frequency using macroeconomic data as of June 2009. The specification includes monthly averages of the semantic scores and federal funds rate, as well as annualized quarterly log-changes in nonfarm payroll employment and in the core PCE

[42]We do not have intra-day estimates of constant-maturity corporate yields.

deflator. Finally the VAR includes monthly averages of par-yields on constant maturity Treasury yields. These yields are estimated from a Svensson-Nelson-Siegel yield curve by Gurkaynak, Sack, and Wright [2007] using off-the-run Treasuries, and the data are available on the website of the Federal Reserve Board.

4 High-frequency interest rate response regressions

In this Section we study the high-frequency response of Treasury (Table 4) and Eurodollar futures (Table 5) rates to changes in the content of FOMC statements, as measured by the linguistic scores. We start by describing the empirical specification and then turn to the estimation results.

Empirical specification We use a high-frequency identification approach to isolate the effects on interest rates of monetary policy actions and communication from other same-day news or events.[43] The model specifications for the Treasury and Eurodollar regressions are analogous, and we only discuss the Treasury one in detail. Let Δy_t^i be the yield change for an on-the-run security with maturity $i = 1, ..., m$ during the tight time window around the FOMC announcement at date t. Under rational expectations, interest rates should only respond to the unanticipated component of target rate decisions and communication during this temporal window. Let MP_t denote the monetary policy surprise. As discussed in Section 3, MP_t is the component of the target rate decision, which is unexpected by market participants as implied by futures quotes.

For the sake of comparison, we separately consider alternative measures of changes in communication, including the change in the human-generated heuristic score, ΔHI, and changes in the four automated measures of semantic orientation either defined on the "hawkish-dovish" antonymy on Google and Factiva, ΔGSO^h and ΔFSO^h, or defined on the larger set of antonymies on the two data sources—ΔGSO^e and ΔFSO^e. Changes in Factiva scores are based on measures recorder after and before a given announcement on time windows of about 1-1/2 days (equation 7). These changes should therefore measure the unexpected change in the content of the announcements along our metric. Changes in the Google and heuristic scores at t are, instead, expressed as differences between the scores at t and $t-1$; see equation 5. These scores are therefore more likely capturing both anticipated and unanticipated components of communication, likely adding some measurement error in the regressions. Under reasonable expectation assumptions, this form of measurement error should bias the estimated coefficients toward zero.

Our empirical setup is designed to allow for cross-equation restrictions, and tests on the vector of the coefficients β^i's measuring interest-rate sensitivities to the linguistic scores and monetary policy surprise.[44] Define $\Delta X_t = [1 \;\; MP_t \;\; \Delta\text{Score}_t]$. The specification expressed in stacked form can then

[43] The validity of this approach is discussed in Cochrane and Piazzesi [2002] for daily data and in Fleming and Piazzesi [2005] intra-day.

[44] For related work in the literature, see Kuttner [2001], Fleming and Piazzesi [2005], and Gürkaynak, Sack, and Swanson [2005], among others.

be written as:

$$\begin{bmatrix} \Delta y_t^1 \\ \Delta y_t^2 \\ \vdots \\ \Delta y_t^m \end{bmatrix} = \begin{bmatrix} \Delta X_t & 0 & \dots & 0 \\ 0 & \Delta X_t & \dots & 0 \\ & & \vdots & \\ 0 & 0 & \dots & \Delta X_t \end{bmatrix} \begin{bmatrix} \beta^1 \\ \beta^2 \\ \vdots \\ \beta^m \end{bmatrix} + \begin{bmatrix} \epsilon_t^1 \\ \epsilon_t^2 \\ \vdots \\ \epsilon_t^m \end{bmatrix}, \tag{8}$$

which can be interpreted as a system of m seemingly unrelated equations. It is important to note that, due to differences in the coverage of FOMC announcements on news and the Internet, not all observations in (8) are measured with the same degree of precision leading to heteroskedasticity in the error term. We can easily account for such variation in measurement precision both for the Google- and Factiva-based regressions. In specifications employing the Google scores we weigh the observations in (8) by the median number of Google hits across all searches pertinent to statement t. [45] In specifications employing Factiva scores we weigh by the total number of sentences in news article relevant for the statement t. [46]

Treasury yields Table 4 reports regression results for Treasury yields. The first column of the Table reports estimates for a benchmark regression only containing the monetary policy surprise MP, while the remaining columns report parameter estimates of models that include the different linguistic scores. The model specification in (8) for Treasury yields is a system of 6 equations having as left-hand-side variable, yields on 3- and 6-month bills as well as those on 2-, 5-, 10-, 30-year coupon bearing Treasuries. The horizontal panels in the Table, report model estimates across yields. First consider the benchmark specification, which only includes a constant and the monetary policy surprise MP (first column of Table 4) and is identical, up to sample coverage, to those considered in earlier literature.[47] Similarly to previous findings, we find evidence of a statistically significant effect of MP only on short-term yields with a substantial drop in the fraction of the variance explained for longer-dated yields.[48]

[45] Across all search units and all sentences. Google weights are constructed from marginal frequencies of search units (i.e. hits(x) using Section 2 notation).

[46] Unweighted results present a loss in precision that occasionally reduces significance below standard confidence levels, but do not affect the qualitative results in high or low frequency. Results available from the authors upon request. We also allow for a general within-announcement (t) covariance structure for the error terms across the m equations in the system:

$$\Sigma_t = \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} & \dots & \sigma_{1m,t} \\ \sigma_{21,t} & \sigma_{22,t} & \dots & \sigma_{2m,t} \\ & & \vdots & \\ \sigma_{m1,t} & \sigma_{2m,t} & \dots & \sigma_{mm,t} \end{bmatrix}$$

within a clustered variance-covariance matrix:

$$V = \begin{bmatrix} \Sigma_1 & 0 & \dots & 0 \\ 0 & \Sigma_2 & \dots & 0 \\ & & \vdots & \\ 0 & 0 & \dots & \Sigma_T \end{bmatrix}.$$

This variance-covariance matrix allows us to obtain standard errors that are robust for general time-varying and within-meeting correlation across error terms.

[47] See Table 1 of Fleming and Piazzesi [2005] and of Gürkaynak, Sack, and Swanson [2005].

[48] The point estimates imply that a 1 standard deviation increase in the monetary policy surprise (about 10 basis points, Table 2) implies positive and significant increases of about 3.5 basis points on yields of 3- and 6-month bills,

The remaining columns of Table 4 report regression results of models that include the different linguistic scores. In each of these models, the coefficient on Δ Score measures the interest rate response in basis points of a unit standard deviation increase in each score (all Δ Scores are standardized to have a zero mean and a unit standard deviation). Column 2 reports regression results for the heuristic score, HI, defined in (1). As shown in the Table the change in the HI score does not appear to contain explanatory power for Treasury yields on the full sample. Much of the explanatory power is lost starting in mid-2007, likely owing to the rigidity of the rule (1) in adapting to changes in the structure of the FOMC statement in the most recent period.[49]

The semantic orientation score included in the regression specification in column 3, GSO^h, is based on the "hawkish-dovish" antonymy using data from Google. According to the parameter estimates, a unit standard deviation increase in the score leads to a hump-shaped response of the Treasury curve around the FOMC announcement. Yield responses to ΔGSO^h are about zero for maturities under a year, peak in the 2- to 5- year sectors of the curve at about two basis points per standard deviation, and are not statistically different from zero for the 30-year bond. The responses are significant at conventional statistical levels in the 2- to 10-year sectors, and are non negligible in magnitude relative to the variation in the dependent variables (Changes in Treasury yields range between 4.5 and 6.5 basis points over the sample, Table 2).[50] As we will discuss in Section 5, yield response are significantly larger after the initial response and reach their maxima more than a year after the initial shock in the scores. The introduction of the semantic scores increases the explanatory power of the regression at medium- and long-term maturities relative to the benchmark regression with an increase in R^2 of about 10 percentage points.[51] As shown at the bottom panel of Table 4, the p-values for the Wald test of equality between the 3-month and 2-year score-coefficients,

and of about 2.5 basis points in the 2-year yield. Note the non-monotonicity in the R^2 reported for the 30−year yield, which presents a higher R^2 than the 2- and 5− year yields. The MP_t coefficient also has the wrong sign for this maturity. This is mostly owes to the inclusion in the sample of the unscheduled meetings of January 3, April 4, and September 17, 2001. The results in terms of the linguistic scores below are robust to the exclusion of these three observations.

[49]More in detail, when estimating the same regression using data through August 2007, we found some predictive power of the HI score in the 6-month to 5-year sector of the Treasury curve. With the onset of the financial turmoil in August 2007 the statistical and economic significance of the HI score has declined significantly, likely owing in part to the notable changes in the structure of the statements—for example, in terms of the number of sentences included in the statement—as compared with those released prior to August 2007. The simple rule defined in (1) to construct the HI score, and particularly its definition at the sentence level, appears rigid in capturing the changing structure of the statement over the entire sample. As we discuss next, the automated semantic orientation scores, instead, appear to adapt more easily to the changing structure of the statement.

[50]We also performed our analysis with the on-the-run 5-year and 10-year Treasury Inflation Protected Securities (TIPS) and the corresponding inflation break-evens. We found positive effects of the statement stance on the real rates, but the estimates were relatively imprecise. The effects on inflation compensation was statistically and economically close to zero. Due to the relative illiquidity of TIPS relative to their nominal counterparts, high-frequency movements in breakevens are probably not very representative of underlying changes in inflation expectations.

[51]Other papers have also shown how FOMC statements and minutes correlate with long-term yields' reactions. Among others this result is confirmed in Gürkaynak, Sack, and Swanson [2005] and Boukus and Rosenberg [2006]. The advantage of our approach is that our scores allow us to precisely identify and quantify the dimension along which the announcement matters. This intuition is lost when employing factor analysis or latent semantic analysis, since the latent factors lack a clear interpretation. Nonetheless such papers deserve credit for pointing at the potential role of FOMC announcements.

as well as between the 3-month and 2-year score-coefficients, both reject statistically confirming the hump-shapedness of the response across maturities. Column 5 reports regression results for the semantic score defined on the "hawkish-dovish" antonymy on Factiva, FSO^h. As shown at the bottom of the Table only 58 observations are included in the estimation, owing to a lack of matches in the Factiva corpus in the sample up to the end of 2003. Although the point estimates are slightly smaller than those obtained using the GSO^h, the results are very similar.

We now turn to parameter estimates for the scores defined on the expanded set of antonymies. Column 4 of Table 4 reports the results for the Google score on the enlarged set of antonymies, GSO^e. Likely, due to the significant error in measuring this score in Google, as discussed above, the coefficients on the scores across maturities are all statistically insignificant (and also enter with a wrong sign). The last column of the Table reports estimates using the Factiva score based on the large set of antonymies, FSO^e. This measure is more precisely estimated than its Google counterpart as we can calculate the empirical frequencies directly on the texts that form the Factiva corpus. Furthermore, we can distinguish the pertinence of the sentences underlying the measure to future, rather than current, policy moves. As for the scores based on the "hawkish-dovish" antonymy, yield responses are hump-shaped across maturities, although the peak of the response is somewhat more pronounced in the 2-year sector for the FSO^e score.[52,53] In sum, with the exclusion of model that includes the Google score based on the expanded set of antonymies, we find that while target rate decisions mainly affect the short-end of the Treasury curve, policy communication have their largest effects on medium-term yields.

Eurodollar futures rates We now turn to a discussion of the response of rates implied by short- medium-dated Eurodollar futures. As discussed in Section 3, these futures, which are among the world's most liquid financial derivatives, settle on realized 3-month Libor rates at specified dates. By studying the response of interest rate futures, we attempt to more precisely characterize yield responses across maturities in an analysis of forward-, rather than spot-, rates in the dollar denominated "risk-free" yield curve.[54] In our analysis, we concentrate on the first 8 quarterly

[52] According to the point estimates, a one standard deviation increase in the FSO^e score, implies an increase in the 2-year yield of about 2.5 basis points, of about 1 basis points for the 6-month bill, and of about 1.5 basis points for the 5-year note. Yield responses at shorter or longer maturities are smaller in magnitude and not statistically significant at conventional levels. The p-values for the Wald tests of equality between the 2- and 30-year coefficients, and between the 3-month and 2-year yields reject the nulls, statistically confirming the presence of a hump in the response of yields across maturities.

[53] It is interesting to note that the size of the coefficients on the monetary policy surprise, MP, remain considerably stable across specifications confirming the quantitative estimates in column (1). This result likely owes to the relatively low correlation between changes in the semantic scores and the MP, as reported in Table 3.

[54] Although the difference, or basis, between Libor and Treasury rates of comparable maturities (also known as the TED spread) were rather stable for most of the sample, the spread between the two rates widened considerably with the start of the financial turmoil in August 2007. The spreads between the two rates has also risen on relatively rare events before 2007, for example ahead of year end in 1999 (Y2K) and at other year ends. For additional discussion about the use of Libor versus federal funds rates in computing market based monetary policy expectations, see Gürkaynak, Sack, and Swanson [2007]. Due to the volatility in this spread, Eurodollar rates likely measured with error "risk-free" rates, implying noisier estimates in our regression. Nonetheless, we focus on Eurodollar futures implied rates due their very high liquidity and our focus on high-frequency responses. In addition, previous literature

contracts, which are the most liquid, therefore covering the 3-month to 2-year sectors of the curve, and for brevity we omit regression results for the third, fifth and seventh contracts in the Tables.

The first column of Table 5 presents results for the benchmark which excludes the linguistic measures. Although the coefficient on MP is positive at all maturities, the size of the coefficients decline for longer-dated contracts, confirming the Treasury-yield findings. Columns 2 through 6 report regression results when including the linguistic measures. As for the case of Treasury yield regressions, we find interest rate responses to the HI and the GSO^e scores not to be statistically different from zero across all maturities. Turning to the scores defined on the "hawkish-dovish" antonymy, the responses to the GSO^h score are increasing under the one year maturity, and about flat for contracts between the fourth and the eighth, or about 1 to 2 years out, with a coefficient of about 3 basis point per standard deviation. The responses to the FSO^h score, which is only defined for 58 observation, are broadly similar, although the point estimates imply responses about 1 basis points smaller and somewhat lower responses at longer maturities. Similarly, the coefficients on the FSO^e score imply a peak response of about 2 basis points for the fourth contract and somewhat smaller responses further out. At the bottom of Table 5, the Wald tests for equality between the coefficient on the first and the eighth futures rate can never reject the null for specifications that include the automated scores.

In sum, although the point estimates are not sufficiently precisely estimated to tightly pin down the peak response of futures rates to FOMC announcements the estimates point to a peak about 1 to 1-1/2 years out. This finding provides some support for the use of Eurodollar implied futures rates at these maturities as proxies of central bank communication in earlier literature, such as Gürkaynak, Sack, and Swanson [2005]. Nonetheless, although the inclusion of the semantic scores leads to a significant increment in the fraction of the variance especially at the one- and two-year maturities, the R^2 of the regression range at about 25 percent at the one-year maturity. This points to some difference in the information captured by variations in the automated scores from those which would be indirectly inferred in using futures rates at these maturities as proxies for changes in communication. We return to this topic in the discussion of the univariate interest rate forecasting regressions in the next Section and in the conclusions.

5 Low-frequency results

This Section studies the link between central bank communication and target rate decisions with measures of inflation, economic activity and nominal risk-free rates. We study these relations using three empirical models: (i) a univariate model to predict short-term interest rates using the semantic scores, (ii) a univariate Taylor [1993]-type forward looking specification for the federal funds rate, and (iii) a vector autoregression (VAR) specification.

attempting to measure the effects of central bank communication has used Eurodollar futures rates at about 1-year maturities as indirect measures of central bank communication

For brevity in this Section we only discuss results for the FSO^e score, which is defined on the expanded set of antonymies using data from Factiva. Overall, we find that this measures performs best in low frequency, especially in differences, as compared with the GSO^e score, defined on Google, and the scores based on the hawkish/dovish antonymy. As previously discussed, because we cannot directly access the text of the webpages, the GSO^e score is likely affected by significant measurement error. In addition, the "hawkish-dovish" scores are also estimated imprecisely until 2003 due to a very limited number of matches on the "hawkish-dovish" antonymy, allowing an analysis on a very limited sample size. For example, the GSO^h measure is based on an average of only about 5 hits for either word in the "hawkish-dovish" antonymy through the end of 2003 compared with about 300 hits on average afterwards.[55] Similarly, due to a lack of hits on the word-pair "hawkish-dovish" in the Factiva corpus, the FSO^h is missing for about half the observations prior to 2003.[56]

5.1 Forecasting short-term interest rates

Empirical specification We study the in-sample predictive power for future realized short-term interest rates of the semantic orientation scores. In particular, we assess whether more (less) hawkish FOMC statements—as measured by the semantic scores—predict (in a Granger sense) higher (lower) short-term rates, as postulated in previous Sections.

The model specification that we consider controls for the unexpected component of the target rate decision, as measured by the monetary surprise, MP, and the information available to market participants ahead of the FOMC announcement as implied by fed funds futures quotes taken 15-minutes before the FOMC announcements. We estimate model specifications at different forecast horizons that correspond to the contract characteristics of the federal funds futures included in the model. In addition, because the liquidity of these futures declines sharply for expiration dates beyond the first few months, we also consider Eurodollar futures on longer horizons.[57] The model specification is a modified Mincer-Zarnowitz regression:

$$(\tilde{r}_{t+\tau_n} - r_{t^-}) = \beta_0^n + \beta_1^n (f_{t^-}^{\tau_n} - r_{t^-}) + \beta_2^n MP_t + \beta_3^n \Delta FSO_t^e + \gamma X_{t^-} + \epsilon_t^n, \tag{9}$$

where t denotes the time of an FOMC meeting. Specifications at each different forecasting horizons τ_n include futures implied rates ahead of the FOMC announcements, $f_{t^-}^{\tau_n}$, and the rate, $\tilde{r}_{t+\tau_n}$, upon which each futures contract settles as a left-hand side. For regressions that include the n^{th} fed funds futures rate, $\tilde{r}_{t+\tau_n}$ is the average federal funds rate on the $n-1$ calendar month following

[55] More precisely, using the notation in () these hits refer to the average across statements of the median hits(dovish) + hits(hawkish) within all search units x in each statement.

[56] In a previous draft of the paper we imputed missing observations in the earlier part of the sample using one-step ahead forecasts from an AR(1) model with the HI score as an exogenous variable using a Kalman filter. In this draft we do not rely on such procedure.

[57] For example, the average notional open interest on fed funds futures in 2008:Q4 was about \$45bn compared to a notional interest of about \$850bn for the fourth (quarterly) Eurodollar futures contract. The open interest for fed funds futures in the earlier part of our sample was even smaller.

the day of the FOMC meeting t. We consider regressions including the second to sixth fed funds futures, corresponding to forecast horizons of up to 5-months. In the case of the n^{th} Eurodollar futures contracts the interest rate $\tilde{r}_{t+\tau_n}$ is the 3-month Libor rate on the settlement date, which occurs in the middle of the settlement month. Because we consider the first eight contracts in the quarterly cycle, the settlement dates fall in the months of March, June, September, and December in either the current or the calendar year following the FOMC statement. The forecasting horizon of the regressions including Eurodollar futures rates therefore range, on average, between 1-1/2 months out (first Eurodollar contract) and 1 year and 11-1/2 months out (eighth contract). For consistency we maintain the weighting procedure described in Section 4 when estimating ().[58]

As it is well known in the literature, the yields included in our regression are highly persistent variables. In order to reduce such persistence, we follow work testing the expectation hypothesis of the term structure of interest rates (Fama and Bliss [1987] and Campbell and Shiller [1991] among others), by subtracting the current level of the intended fed funds rate and Libor rates, respectively, from the realized rates, $\tilde{r}_{t+\tau_n}$ and the futures rate $f_{t^-}^{\tau_n}$ in the fed funds and Eurodollar futures regressions. The dependent variable in (58) is therefore the difference between realized spot and current rates, whereas the independent variable is the futures-spot spread.[59]

Under forecasting efficiency of futures implied rates only the new information contained in the FOMC statement, or the target rate decision, can have additional predictive power for realized rates around the FOMC announcement. We capture the unexpected component of the target rate decision using MP, and the new information in the statement using the change in the FSO^e score, which, as defined in (7), uses information from the Factiva-corpus right before and after the release of the FOMC announcement. The change in the semantic score, are standardized to a have a unit standard deviation, and the yield variables included in the regressions are expressed in basis points. The coefficient β_3^n therefore measures the basis point response in the realized short-term rates to a unit standard deviation unexpected increase in the hawkishness of the FOMC announcement as implied by the reference text in Factiva.

It is important to note, however, that although we cannot reject the null of forecasting efficiency of futures rates in a regression that only includes the futures-spot spread and a constant (that is, we cannot reject the joint condition $\beta_1^n = 1$ and $\beta_0^n = 0$), consistent with the findings of Gürkaynak, Sack, and Swanson [2007], work of Piazzesi and Swanson [2008] find excess returns on fed funds and Eurodollar futures to have been strongly countercyclical between 1988 and 2003, a finding that they attribute to the presence of time-varying risk premia in market prices. Following their work we also consider a regression specification that includes the additional set of controls, X_{t^-},

[58] In using these weights, we place more emphasis on observations that are more precisely measured. Refer to Section 4 for detail on the weighting scheme. We find a limited impact of these weights on either point estimates or standard errors of ().

[59] Futures rates are equal to forward rates up to some convexity adjustments because of marking-to-market on exchanges of these contracts. These convexity adjustments are very small at the maturities that we consider, and we therefore treat futures implied rates as forward rates in what follows (Burghardt [2003]).

measured ahead of the FOMC announcement, which attempt to proxy for the time variation in term premia in the futures rates. The set of controls includes the 10- to 2-year slope of the Treasury yield curve, the credit spread between 10-year BBB-rated corporate bonds and Treasuries, and a dummy variable for NBER-dated recessions.[60] Finally, because of the overlapping forecasting horizons in (58) and the corresponding moving average component in the error term, as well as the varying forecasting horizon due to the uneven distribution of FOMC dates over the course of the calendar year, we compute Newey and West [1987] heteroskedasticity and autocorrelation consistent covariance matrices with truncation lags equal to 1-1/2 times the forecasting horizon, τ_n. We follow Cochrane and Piazzesi [2005] in setting a longer truncation lag than the number of overlapping observations to counteract the under-weighting of distant covariance terms implied by the Newey-West kernel weighting function. Finally, although the Newey-West adjustment accounts for the time-series properties of the error term, it is important to point out that the elements of the variance covariance matrices at the longest forecasting horizons are likely imprecisely estimated, due to the significant length of the truncation lags relative to the estimation sample. For the same reason, the point estimates at distant horizons are probably estimated with less precision than implied by the Newey West correction. Bearing these caveats in mind, we turn next to the regression results.

Estimation results The parameter estimates of the forecasting model for the federal funds and 3-month Libor rates are reported in Tables 6 and 7, respectively. The upper-panel of Table 6 reports estimates for the federal funds rate models. The different columns in the Table correspond to specifications including the second to sixth fed funds futures rate, corresponding to a forecast horizon of 1- to 5-months. As it can be seen from the different columns, the sensitivity of future rates to unit standard deviation increase in the FSO^e build up monotonically with the forecast horizon, ranging from about 5 basis points 1-month out, to about 25 basis points at the 5-month horizon.[61] The bottom-panel of Table 6 repeats the same regression exercise when controlling for time varying risk-premia using an NBER recession dummy, the slope of the Treasury curve and a credit spread. The regression coefficients on ΔFSO^e decline slightly and range between 3 and 20 basis points per standard deviation in ΔFSO^e in these specifications.

The regression results for the Libor rate forecasting model that exclude the term-premium controls are shown in the upper-panel of Table 7. The columns report results for the models including the first to eighth Eurodollar futures rate, corresponding to 3-month realized Libor starting about 1-1/2 to about 2 years out. As for the federal funds rate forecasting regression the coefficients on ΔFSO^e are monotonically increasing in magnitude and range between 4 and 70 basis points per standard

[60] We have also included the real-time 12-month change in nonfarm payroll employment used in Piazzesi and Swanson [2008], but found little predictive power of this variable in our shorter sample.

[61] In terms of the other controls included in the regression, the unexpected target decision MP contains predictive power for realized rates only at very short horizons. Although the regression coefficient on the futures-spot spread, $(f_{t^-}^{\tau_n} - r_{t^-})$, is somewhat larger than one, we cannot reject the null of efficiency for these rates.

deviation in ΔFSO^e.[62] With the exception of the first few maturities the coefficients on the futures-spot spread are well above unity pointing to the lack of efficiency of the Eurodollar implied rates, although based on the standard errors we cannot reject the null of forecasting efficiency. The bottom panel of the Table, reports parameter estimates when the proxies for the time-varying term premia are included in the regression. Following the inclusion of these controls, the coefficients on the futures-spot spread decline sharply, and they now only enter significantly for the first few contracts, highlighting the lack of efficiency and presence of term premia in these rates. Importantly for our analysis, the coefficients on ΔFSO^e, also decline notably after including the term premium proxies, and now range between about 1 basis point and 45 basis points, with a peak coefficient of about 55 basis points at a forecast horizon of about 1-2/3 years (seventh Eurodollar contract).

As noted in Section 4, although changes in Eurodollar futures implied rates around FOMC announcements exhibit significant correlation with the semantic scores, when the change is conditioned on the monetary policy surprise, MP, such correlation is fairly limited. As a final exercise we re-estimate the model (58) using quotes taken after, rather than before the FOMC announcements, to construct the futures-spot controls. In other words, we substitute $f_{t^+}^{\tau_n}$ to $f_{t^-}^{\tau_n}$ in (58), where t^+ is 20 minutes after the FOMC announcement at date t. The results of this analysis are shown in Tables 8 and 9, respectively, for the the federal funds and 3-month Libor rate models. Comparing the regression coefficients on ΔFSO^e with those in Tables 6 and 7, it is apparent that the point estimates are little affected by the inclusion of futures quotes taken after, rather than before, the release of the FOMC statement. Of course, one would expect futures quotes to contain all relevant information to predict realized rates. However, based on the in-sample analysis, implied rates do not appear to fully achieve such task, perhaps due to the presence of term premia. These results appear to indicate that the common use by practitioners and in previous research literature (for example, Gürkaynak, Sack, and Swanson [2007]) of futures quotes to indirectly measure the content of central bank communication could be somewhat limited, as these quotes might not necessarily reflect all the information contained in the statements.

In conclusion, we find significant predictive content of changes in the semantic score FSO^e, both in the federal funds and Libor rate forecasting regressions, with a 1 standard deviation increase in the FSO^e score associated with a hump-shaped response of realized rates having a peak about 1-2/3 years out (model including term-premium controls). We also find that the information of the FSO^e does not appear to be fully incorporated in futures quotes following the FOMC announcement. We will find a similar predictive performance of the FSO^e score in Section 5.3, which considers a multivariate (VAR) model specification that includes the FSO^e scores and federal funds rates, as well as measures of inflation and economic activity. Before turning to that model, we interpret the information in the FOMC statement and semantic score through the lens of a forward-looking Taylor rule model.

[62] The unexpected policy rate decision once again only contains predictive power for very short horizons.

5.2 Taylor rule and automated scores

This Section's goal is to assess the type of information contained in FOMC statements as measured by the automated semantic scores. We first estimate the parameters of a forward looking Taylor rule model for the federal funds rate. Based on these estimates, we decompose realized values of policy rates into two orthogonal components: 1) A systematic component, or "Taylor rule rate", which is the portion of the policy rate decision explained by forecasts of inflation and of output gap, and 2) a residual component, which we refer to as the "interest rate gap".[63] We then compute cross-correlations between the two components and the semantic score at different leads and lags, and study whether the automated scores hold a stronger correlation with either of the two components, as well as the score's leading properties. We then attempt to interpret the information contained in the statements based on these correlation measures.

We consider a Taylor-rule specification that incorporates partial interest rate adjustment, to account for the observed sluggishness of policy rates (Clarida, Galí, and Gertler [2000]). In addition, the rule assumes a forward looking behavior of the central bank, by including forecasts, rather than current realized measures of inflation and output gap, as determinants of interest rate decisions. As shown by Orphanides [2001], real-time policy recommendations can differ substantially from those obtained using revised data. We therefore attempt to match the information set of FOMC members as closely as possible by using the Greenbook forecasts for the GDP deflator and the output gap. These forecasts are prepared by the staff of the Federal Reserve Board ahead of each scheduled FOMC meeting. The Greenbook forecasts are unfortunately only available to the public with a 5-year lag. We supplement the Greenbook forecasts starting in 2003 with forecasts of inflation and real GDP from the Survey of Professional Forecasters (SPF), and of potential GDP from the Congressional Budget Office (CBO). It is unlikely that these forecasts were a direct input of actual policy decisions by FOMC members. However, because they were formed around the same time in which the Greenbook forecasts were made, they were based on information sets similar to those available to Board staff ahead of FOMC meetings. In this sense we think of them as proxying the true Greenbook forecasts.[64]

Measures of inflation and of the output gap display their most important fluctuations at business cycle frequencies. Because the time span starting in 1999 during which the automated scores are defined, only includes two business cycles (as defined by the NBER), the parameters of the Taylor rule are unlikely to be well identified on this sample. Given our two-step procedure, however, we can decompose policy decisions since 1999 using parameter estimates obtained on a longer sample.

[63] As discussed below the Taylor rule model also includes as explanatory variables the lagged interest rate. This term is not included in what we call "Taylor rule rate".

[64] Additional detail on the data is available in Section 3. We closely follow Orphanides [2003] in using Greenbook and SPF forecasts at mid-quarter. Prior to 1992, output concept to measure economic activity and inflation is GNP and not GDP. Also the unit base year for the GNP/GDP measures change over the sample following those used by the BEA in each. See the Federal Reserve Bank of Philadelphia web-pages on "Real-time forecasts" for additional detail. The CBO forecasts are only available twice a year, generally in the month February and August. The federal funds rate is a quarterly average.

Because monetary policy decisions under the Volcker chairmanship were arguably dominated by idiosyncratic factors, we estimate parameters of the Taylor rule model between September 1987 and December 2008, a period that covers the tenures of chairmen Greenspan and Bernanke. As for the univariate forecasting regressions, we again focus the analysis on the FSO^e score, which is defined on the extended set of antonymies using the Factiva corpus of text. The model specification that we use closely follows Orphanides [2003]:

$$i_t = \alpha \ i_{t-1} + \beta_0 + \beta_\pi \ \pi^a_{t+3} + \beta_\pi \ \Delta y^a_{t+3} + \beta_y \ y_{t-1} + \epsilon_t, \tag{10}$$

where i_t is the federal funds rate (quarterly average). The variable π^a_{t+3} is the 4-quarter inflation forecast starting in $t-1$. Because the BEA releases preliminary GDP estimates with a delay of about 1-1/2 months, quarter $t-1$ represents the one for which the most recent data are available when the forecasts are formed at t. The variable y_{t-1} denotes the output gap at time $t-1$ and $\Delta y^a_{t+3} = y_{t+3} - y_{t-1}$, is the forecast of the 4-quarter change in the output gap starting in quarter $t-1$. The "interest rate gap" is equal to the residual ϵ_t, and we define the "Taylor rule rate" as the predicted component of the rule with the exclusion of the lagged interest rate term:

$$i^T_t \equiv \beta_0 + \beta_\pi \ \pi^a_{t+3} + \beta_y \ \Delta y^a_{t+3} + \beta_y \ y_{t-1}.$$

The rule in (10) is more general than the one originally considered by Clarida, Galí, and Gertler [2000], unless the condition $\beta_\pi = \beta_y$ is satisfied. Linear least square estimation of the parameters in (10) yields to:

$$\hat{i} = .85_{[.04]} \ i_{t-1} - .11_{[.12]} + .39_{[.09]} \ \pi^a_{t+3} + .42_{[.06]} \ \Delta y^a_{t+3} + .22_{[.03]} \ y_{t-1}.$$

where the subscripts in square brackets denote Newey-West standard errors with a 4-quarter truncation lag rule. The point estimates are consistent with those obtained by Orphanides [2003] among others. The interest rate decision is characterized by a significant degree of inertia, and the point estimates on the inflation term guarantee stability in economies in which interest rates are set using these forwarding looking rules.[65] In addition, β_π is significantly different from β_y, supporting the use of the more general specification that includes the forecast of the growth rate of the output gap.

After obtaining estimates for ϵ_t and i^T_t, we calculate the cross-correlation functions $\text{Corr}(\hat{\epsilon}_t, FSO_{t-J})$ and $\text{Corr}(\hat{i}^T_t, FSO_{t-J})$, which are shown in Figures 4 and 5, respectively. As it can be seen from the two figures, the semantic score, FSO^e, displays rather strong correlations with both components (about 70 percent). In other words, FOMC announcements contain information regarding policy decisions based on changes in the outlook for inflation and economic activity ("Taylor rule rate"), as well as other factors included in the "interest rate gap". These factors are orthogonal to the

[65] More precisely, the price level is stable when the nominal interest rate, i, increases more than proportionally with an increase in inflation (Clarida, Galí, and Gertler [2000]). The point estimates satisfy these condition as $\beta_\pi/(1-\alpha) = 2.6$.

systematic component by construction. The residual in the Taylor rule is often interpreted in the literature as a shock to the preference of policy makers. This could be due to an imperfect observability of the state of the economy in real-time (Orphanides and Williams [2007]), or alternatively, the weights given by the FOMC to its dual mandate of price stability and economic growth could shift over time (for example, Boivin, Dong, and Ang [2008]). Alternatively, the residual can be interpreted as omitted variables that affect policy decisions and are uncorrelated with the systematic component. For example, Gertler and Karadi [2009] and Taylor [2008] have proposed rules according to which central banks respond to financial stress indicators to achieve monetary policy objectives at times of limited pass-through of target policy rates on other yields. Although a casual reading of the FOMC statements over the past 10-years seems to provide some support to both interpretations, our semantic scores cannot distinguish among these alternatives.

Irrespective of the underlying determinant of the Taylor residuals, however, the cross-correlation functions highlight another interesting feature of the linguistic measure. As it can be seen from Figures 4 and 5, whereas the FSO^e displays the highest correlation with the systematic component of policy, i_t^T, contemporaneously, the score leads the residual by about 1- to 2-quarters. Although, our analysis does not identify reasons for policy deviations from the prescriptions of a "canonical" forward-looking Taylor rule, the results suggest that the FOMC communicates information regarding such deviations with significant lead before the actual policy rate decisions are taken. Because longer-term yields (see Section 4) and other macroeconomic variables could respond to such information, standard empirical monetary models considered in the literature (for example, Christiano, Eichenbaum, and Evans [1999]) might incorrectly identify measures of monetary policy shocks. The next section considers a small-scale VAR that includes a measure of communication.

5.3 Vector autoregression analysis

So far we have analyzed in univariate model specifications interest rate responses to shocks to the semantic scores (Section 4), the predictive power of the scores for future realized rates (Section 5.1), and the information content of scores as implied by a Taylor rule (Section 5.2). In this Section, we revisit some of these results within a VAR multivariate framework. This model identifies policy shocks with a very different approach and thus validates some results of the previous sections. In addition, based on these identification scheme, we can study interest rate responses to shocks in the following months , and not just during the tight temporal windows around FOMC announcements as in Section 4.

The VAR models include two monetary policy instruments: a policy interest rate (the federal funds rate) and the Fed's communication about future movements in the policy rate in the FOMC statements. After presenting the model and the identification for the monetary policy shocks, we turn to a discussion of the impulse responses and the forecast error variance decompositions of interest rates and real variables to unexpected innovations in the two instruments.

Empirical specification Our sample of analysis begins in May 1999 and ends in December 2008. In order to capture the response of Treasury yields at different maturities to the same exact policy shocks, while keeping the number of parameters low, we estimate six VAR models that feature identical measures of inflation and economic activity, the federal funds rate and the FSO^e score. The specifications differ for the maturity of the nominal yield included. More precisely, let $\mathbf{Y}_t^i = [\mathbf{X}_t, \mathbf{S}_t, R_t^i]'$ denote the vector of variables in the VAR model i : $\mathbf{X}_t$ includes the 3-month core-PCE inflation rate and the 3-month change in non-farm payroll employment; both variables are reported as basis point changes expressed at an annual rate. The vector $\mathbf{S}_t$ denotes the policy block composed of, in order, the semantic orientation score and the federal funds rate. Finally, R_t^i, denote Treasury par-yields with maturities ranging between 3-months and 30-years. We refer to the variables $\mathbf{Z}_t = [\mathbf{X}_t, \mathbf{S}_t]'$ as the "core" variables of the models.

The VAR models identify monetary policy shocks using a commonly used recursiveness assumption: First, the change in the aggregate price level and in non-farm payroll employment, $\mathbf{X}_t$, do not respond contemporaneously to innovations in the policy block $\mathbf{S}_t$ and the yield R_t^i. Furthermore, within the policy block, $\mathbf{S}_t$, the semantic score is ordered first, so that the federal funds rate responds immediately to innovations in the score. We find this ordering to be preferable to the alternative of the federal funds rate being ordered first in $\mathbf{S}_t$. Following the discussion of Section 5.2, we interpret the FSO^e as partially indicating a change in the preferences of policy makers, such as a more hawkish stance towards inflation. The federal funds rate can respond immediately to the FSO^e when it is ordered first in the VAR. Nevertheless, qualitatively the main findings in this Section do not critically depend on the ordering of the two policy instruments within $\mathbf{S}_t$.

Most of the VAR models considered in earlier literature to identify monetary policy shocks do not include information regarding the term structure of nominal yields. In this respect, our model is closely related to this literature in that we assume that innovations in the yields R_t^i's do not affect any of the core variables, $\mathbf{Z}_t$, neither contemporaneously nor with a lag. Instead, the yield R_t^i can respond contemporaneously to innovations in the core variables.[66] The structural form of the VAR models can be written as:

$$\mathbf{a}\,[\mathbf{X}_t, \mathbf{S}_t, R_t^i]' = \mathbf{A}\,[\mathbf{X}_t, \mathbf{S}_t, R_t^i]' + \sigma\,[\epsilon_t^{\mathbf{X}}, \epsilon_t^{\mathbf{S}}, \epsilon_t^{R^i}]', \tag{11}$$

for $i = \{$3-month, 6-month, 1-year, 3-year, 10-year, 30-year$\}$, where:

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} & 0 \\ \mathbf{a}_{21} & \mathbf{a}_{22} & 0 \\ \mathbf{a}_{31} & \mathbf{a}_{32} & 1 \end{bmatrix}, \ \mathbf{A}(L) = \begin{bmatrix} \mathbf{A}_{11}(L) & \mathbf{A}_{11}(L) & 0 \\ \mathbf{A}_{21}(L) & \mathbf{A}_{22}(L) & 0 \\ \mathbf{A}_{31}(L) & \mathbf{A}_{32}(L) & A_{33}(L) \end{bmatrix}, \tag{12}$$

and the matrix σ is diagonal. The diagonal terms in the matrices $\mathbf{a}_{11}$ and $\mathbf{a}_{22}$ of (12) are equal to

[66] For a review of this literature, see Christiano, Eichenbaum, and Evans [1999]. The specification of the model that we consider closely resembles that of Evans and Marshall [1998]. For a model, which, instead, uses information on asset yields to identify monetary policy shocks see Piazzesi [2005].

one, and the innovations $\epsilon_t^{\mathbf{X}}$, $\epsilon_t^{\mathbf{S}}$ and $\epsilon_t^{R^i}$ in (11) are the serially- and mutually-uncorrelated i.i.d. structural shocks. It is important to note that, because of the zero elements in (12), the structural shocks of the policy block $\epsilon_t^{\mathbf{S}}$ do not depend on the Treasury yield included in each VAR model. In other words, each VAR model considered identifies the same monetary policy shocks.[67] Based on the Akaike Information Criterion, we include six lags of the relevant variables in the model specifications in (11), and because of the zero-restrictions in (12), we estimate the parameters in (11) as a seemingly unrelated system.[68]

Results Figures 6 and 7 show responses of the core variables, $\mathbf{Z}_t$, to an unexpected one standard deviation increase in the innovations to the FSO^e score and the federal funds rate $\epsilon_{\mathbf{t}}^{\mathbf{S}}$. Figures 8 and 9 display Treasury yield responses to these shocks. All responses in the figures are absolute deviations in basis points from the unshocked values, with the exception of the FSO^e score, which is reported as an absolute deviation in the (unscaled) units of the variable. In all figures, the shaded areas represent two-standard error bootstrapped confidence bands for the corresponding impulse response.[69] Before discussing the results, it is important to note that a positive innovation to both variables can be interpreted as contractionary monetary policy shocks, whereas positive innovations to the federal funds rate directly feed into higher short term rates, a positive innovation to the linguistic score may affect short term rates only in that it indicates a more "hawkish" stance of policy.

Consider the response of the core variables to an unexpected positive shock to the FSO^e score. As shown in the bottom-right panel of Figure 6 the response of the federal funds rate is hump-shaped with a peak of about 30 basis points about a year after the shock. The response of the FSO^e score (bottom-left panel) to its own innovation peaks almost immediately and then declines monotonically, returning to its pre-shock level in less than a year. The responses of both core inflation and employment growth to a shock in the semantic orientation score are in general not statistically different from zero likely due to the limited sample size (upper-panels). Following a shock to the linguistic score, the core inflation rate falls a few basis points for the first year. The response of nonfarm payroll employment to a positive innovation to the FSO^e score is quite persistent. It is initially positive for a few months and it then turns negative after about 1-year. Figure 7 reports responses of the core variables to an unexpected (positive) innovation in the federal funds rate. As shown in the bottom panel, the response of the FSO^e score is initially slightly positive and then becomes negative after 6-months. This variation is economically small and always remains within 5 percent of the in-sample standard deviation of the FSO^e score. The

[67] Note that, as for the case of the high-frequency analysis of Section 4 the VAR model does not impose absence of arbitrage opportunities across different maturities in calculating the yield responses. Although interesting, these restrictions are beyond the scope of this section.

[68] Ivanov and Kilian [2005] find that, for monetary models of the sort considered here, the AIC provides the most accurate estimate of impulse responses in small sample and data observed at a monthly frequency.

[69] To obtain the confidence bands, we resample 1,000 times from the fitted residuals of (11). The confidence bands are then constructed as the point estimates of the impulse response coefficients plus/minus two standard deviations of the impulse response coefficients across the resampled datasets.

response of the federal funds rate to its own innovation is very persistent: It is positive and then becomes negative after about 2-years. As for the responses to innovations in the FSO^e score, the responses of core-inflation rate and nonfarm payroll employment to the federal funds rate shock are also not statistically significant.[70] Even in the relatively short sample size, however, the model estimates imply that an unexpected positive innovation in the semantic orientation score, that is, an unexpectedly more hawkish FOMC statement, is followed by an increase in the federal funds rate confirming the results found in the univariate model analysis in Section 5.1. The response of short-term interest rates to shocks to the FSO^e score implied by the VAR and univariate model are very similar at the 1-year horizon (about 30 basis points). Nonetheless, we note some differences in the results of the two models, as the response of interest rates declines in the VAR, but increase further for a couple of quarters in the univariate model. Some of these differences are likely attributable to the very different identification schemes used in the two analyses.

We now turn to the forecast error variance decompositions of the core variables, which describe the portion of the conditional k-step ahead forecast error variance of each variable that can be accounted for by the innovation in the semantic orientation score and the federal funds rate. The estimated variance decompositions are shown in Table 10.[71] For each core variable, there are two columns in the Table showing the portion of the variance accounted for by the semantic orientation score—column FSO^e in the Table—and the federal funds rate—column FR. As shown in the second-to-last column, the shock to the linguistic score accounts for a significant portion of the forecast error variance of the federal funds rate with a maximum of about 45 percent 6- and 12-months out. The variance of the federal funds rate accounted for by its own shock, instead, is monotonically decreasing, with a maximum of about 55 percent at 3-months, and only about 25 percent 1-year ahead. Interestingly the variance in the federal funds rate is largely accounted for by the FSO^e score at the 6- and 12-month horizon, indicating that at such horizon previous communication plays a very important role. The variance of the semantic orientation score accounted for by its own shock is monotonically decreasing and is quite large, with a maximum of about 90 percent after three months and 55 percent 1-year ahead. The innovation of the federal funds rate, instead, accounts for a negligible portion of the score variance at all forecasting horizons.[72]

[70]The growth rate of nonfarm payroll is initially positive, and then becomes negative after about 1-year. The core inflation rate fluctuates in the first year following the shock, and then turns slightly negative after the first year. The sample considered in this analysis is relatively short and only includes a few business cycle fluctuations. It should not come as a surprise that the responses of the real variables to shocks to the FSO^e and federal funds rate are imprecisely estimated. Arguably, their shapes do not closely resemble findings in previous related literature that has analyzed a much longer data samples. The data sample analyzed in the literature that uses VAR models to identify monetary policy shocks, typically starts in the early sixties and ends in the late nineties. In the VAR analysis, we obtained similar results when using alternative measures of inflation—for example, headline—or output, such as an index of industrial production. The inclusion of commodity price indices, which are often considered to solve the "price puzzle", also did not significantly affect the shape of the impulse responses. Finally, the responses of core inflation and employment growth to a shock to the federal funds rate were similar when excluding the semantic orientation score from the models.

[71]Bootstrapped standard errors of the point estimates are reported in square brackets below each corresponding value.

[72]The portions of the variances of the core inflation rate and of the nonfarm payroll employment growth rate that is

We now turn to the responses of Treasury yields to unexpected shocks to the semantic orientation score, FSO^e score, and the federal funds rate. As shown in Figure 8, yields at all maturities rise on impact after a positive innovation to the FSO^e score, with magnitudes that generally decline with maturities. The responses of the 3-month and 6-month yields are hump-shaped with a peak response of about 25 basis points 6- to 12-months; both responses are statistically different from zero for more than a year after the shock. The responses of yields at 2- and 5-year maturities display a slight hump, with a peak at about 3-months of about 15 basis points, and are statistically significantly different from zero for somewhat less than 6-months. The responses of yields at the longest maturities are never statistically different from zero. The yield responses to a federal funds rate shock are shown in Figure 9. The responses of the 3- and 6-months yield display a slight hump and are larger in magnitude compared with the responses to a FSO^e shock at short horizons. Instead, all responses at maturities beyond two years are smaller in magnitude than for a FSO^e shock, and are never statically significant. This pattern of more pronounced responses to policy rate shocks at short-horizons and larger responses to FSO^e shocks at longer maturities is analogous to the high-frequency results of Section 4. The portion of the forecast error variance of the yields accounted for by the FSO^e and the fed funds rate shocks are reported in Table 11. The federal funds rate shock account for 30 and 20 percent of the conditional variance, respectively, of the 3- and 6-months yields at the 3-months horizon; the portion of the variance declines at longer horizons for these two yields, and it is always tiny at longer maturities. The FSO^e score shock, instead, accounts for considerable conditional variation for yields with maturities of up to 5-year. For the 3- and 6-months maturities the portion of conditional variance explained is hump-shaped with a peak of roughly 55 percent 6- to 12-months after the initial shock. At the 2- and 5-year maturities the portion of the forecast variance is fairly stable for the first 6-months at about 20 and 10 percent, respectively.

In sum short-, medium-term Treasury yields and the federal funds rate increase after a shock to the FSO^e score. The responses of short-maturity Treasuries and the federal funds rate to these shocks display significant humps that peak at about 1-year after the shock. The responses of medium-term Treasuries are less long lived. The FSO^e score explains about half of the conditional variation of these short-dated rates at the 1-year horizon. With the exception of the immediate responses of short-term Treasury yields, our analysis indicates that central bank communication has been a more important determinant of nominal rates than immediate policy rate decisions. In this sense, words have been more important than actions for Treasury yields over this past decade.

6 Conclusions

In this paper we presented a novel approach to measure the content of central bank communication regarding future policy rate decisions, and applied these measures to FOMC statements starting in

accounted for by either of the two monetary policy shocks are generally less than 5 percent at the different horizons, and the estimates are imprecise.

1999. These new measures rely on a branch of research in computer science and natural language processing that has been rather untapped by economists. The approach of this paper, in particular, has the advantage of being unsupervised, intuitive, and replicable across researchers.

Based on the linguistic measures, in high-frequency data we find that while short-term Treasuries respond to immediate (unexpected) policy rate decisions, longer-dated yields mainly react to changes in the content of communication as measured by our semantic scores. In lower frequency data, we find that changes in communication predict future policy rate actions with a lead of more than a year in univariate interest-rate forecasting models, in which shocks are identified using intra-day futures quotes, and in a VAR model with shocks identified recursively using measures of economic activity and inflation. The empirical results also show how changes in communication have been more important than the contemporaneous setting of policy rates in determining longer-term nominal yields. Based on a forward-looking Taylor rule decomposition, we also find that FOMC announcements contain significant information regarding both the predicted and the residual component of the rule-based interest rate decision. In addition, the semantic scores lead the rule-based policy rate residual—or monetary policy shock—by a few quarters.

By emphasizing the role of central bank communication, rather than the immediate setting of interest rates, our analysis highlights an important dimension of monetary policy that has received only limited attention in the empirical literature on monetary economics. The importance of communication in affecting the term structure of nominal rates, along with the leading properties of the semantic measures relative to rate decisions, suggests that monetary models used in earlier literature—for example, canonical VARs as in Christiano, Eichenbaum, and Evans [1999]—might have omitted an important measure when identifying monetary policy shocks.

Work in the economic and finance literature (such as Gürkaynak, Sack, and Swanson [2005]), as well as central bankers and market practitioners, have generally proxied changes in central bank communication indirectly through changes in market rates. While these measures reflect expectations of market participants, they are only an imperfect measure of such expectations because of movements in term premia. While our semantic measures are quite correlated with futures implied rates, especially at the one- to two-year horizon, the correlation between these measures is far from perfect, and our findings seem to suggest that these quotes have only imperfectly reflected the content of communication (as implied by our measure and our in-sample analysis), at least in predicting future realized policy decisions.

In the paper we present measures of communication based on information from the Internet, obtained via searches in Google, and from news sources included in the Dow Jones Factiva database. While the Internet contains information by several order of magnitudes greater than what is accessible on news, our Google-based semantic measures only rely indirectly on such information in using hit-counts on the search engine, rather than actual matches on the underlying text. Based on our empirical results, and under the current state of technology, we find that having a direct access and control of the information underlying the corpus of text, even if smaller, is very important. Indeed, our empirical findings suggest that the Factiva-based scores are estimated with a significantly greater precision as compared to the Google-based scores.

The information used to construct our semantic measures relies on commentaries by members of the press, market practitioners, and investors. In this sense, our semantic measures could be used as an alternative source of agents' expectations and opinion to market instruments or surveys. The semantic measures used in this paper could be of even more importance in other fields that lack such alternative measures of average opinion. Investigations using the semantic scores could, for example, be applied in the empirical validation of recent theoretical contributions aiming at understanding the role of communication in the interaction between economic agents (such as Morris and Shin [2007]). Applications outside policy announcements may range from the orientation of political campaigns to communication between firms and investors.

References

Antweiler, W., and M. Z. Frank (2004): "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *The Journal of Finance*, 59(3), 1259–1294.

Bernanke, B. S. (2004): "Fedspeak," Speech delivered at the AEA meetings, San Diego, California.

Bernanke, B. S., V. R. Reinhart, and B. P. Sack (2004): "Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment," *Brookings Papers on Economic Activity*, 70(2004-2), 1–100.

Bird, S., and E. Loper (2006): "Natural Language Toolkit," Available at: `nltk.sourceforge.net`.

Blinder, A. S., M. Ehrmann, M. Fratzscher, J. D. Haan, and D.-J. Jansen (2008): "Central bank communication and monetary policy: A survey of theory and evidence," *Journal of Economic Literature*, 46(4), 910–945.

Boivin, J., S. Dong, and A. Ang (2008): "Monetary Policy Shifts and the Term Structure," *mimeo, Columbia University.*

Boukus, E., and J. V. Rosenberg (2006): "The information content of FOMC minutes," mimeo, Federal Reserve Bank of New York.

Brill, E. (1994): "Some advances in transformation-based part of speech tagging," *Proceedings of the twelfth national conference on Artificial intelligence*, 1, 722–727.

Burghardt, G. (2003): *The Eurodollar futures and options handbook*. McGraw-Hill.

Campbell, J., and R. Shiller (1991): "Yield spreads and interest rate movements: A bird's eye view," *Review of Economic Studies*, pp. 495–514.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): "Monetary policy shocks: What have we learned and to what end?," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford. Elsevier, North Holland.

Church, K. W., and P. Hanks (1990): "Word association norms, mutual information, and lexicography," *Computational Linguistics*, 16(1), 22–29.

Clarida, R., J. Galí, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147–180.

Cochrane, J., and M. Piazzesi (2002): "The Fed and Interest Rates: A High-Frequency Identification," *American Economic Review*, 92(2), 90–95.

——— (2005): "Bond Risk Premiums," *American Economic Review*, 95, 138–160.

Demers, E. A., and C. Vega (2008): "Soft Information in Earnings Announcements: News or Noise?," *FRB International Finance Discussion Paper No. 951.*

Eggertsson, G., and M. Woodford (2003): "The Zero Bound on Interest Rates and Optimal Monetary Policy.," *Brookings Papers on Economic Activity*, 1, 139–234.

Evans, C. L., and D. A. Marshall (1998): "Monetary policy and the term structure of nominal interest rates: Evidence and theory," *Carnegie-Rochester Confer. Series on Public Policy*, 49, 53–111.

Fama, E., and R. Bliss (1987): "The information in long-maturity forward rates," *American Economic Review*, pp. 680–692.

Fleming, M., and M. Piazzesi (2005): "Monetary Policy Tick-by-Tick," Discussion paper, University of Chicago, mimeo.

Gentzkow, M., and J. M. Shapiro (2006): "What Drives Media Slant? Evidence from U.S. Daily Newspapers," NBER Working Paper No. 12707.

Gertler, M., and P. Karadi (2009): "A Model of Unconventional Monetary Policy," *mimeo, NYU.*

Gorodnichenko, Y., and M. D. Shapiro (2007): "Monetary policy when potential output is uncertain: Understanding the growth gamble of the 1990s," *Journal of Monetary Economics*, 54(4), 1132–1162.

Gürkaynak, R. S., B. Sack, and E. Swanson (2005): "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*, 1, 55–93.

——— (2007): "Market-based measures of monetary policy expectations," *Journal of Business and Economic Statistics*, 25(2), 201–212.

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

Hatzivassiloglou, V., and K. McKeown (1997): "Predicting the semantic orientation of adjectives," *Proceedings of the 35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL*, pp. 174–181.

Ivanov, V., and L. Kilian (2005): "A Practitioner's Guide to Lag Order Selection for VAR Impulse Response Analysis," *Studies in Nonlinear Dynamics and Econometrics*, 9, Article 2.

Kohn, D. L. (2005): "Central Bank Communication," Speech delivered at the AEA meetings, Philadelphia, Pennsylvania.

Kuttner, K. N. (2001): "Monetary Policy Surprises and Interest Rates: Evidence From the Fed Funds Futures Market," *Journal of Monetary Economics*, 47, 523–544.

Landauer, T. K., and S. Dumais (1997): "A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge," *Psychological Review*, 104(2), 211–240.

Liu, H. (2004): "MontyLingua: An end-to-end natural language processor with common sense.," .

Manning, C., and H. Schütze (1999): *Foundations of statistical natural language processing.* MIT Press.

Mishkin, F. (2007): "Can central bank transparency go too far?," *Monetary policy strategy*, p. 89.

Morris, S., and H. Shin (2007): "Optimal Communication," *Journal of the European Economic Association*, 5(2-3), 594–602.

Newey, W. K., and K. D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Orphanides, A. (2001): "Monetary policy rules based on real-time data," *American Economic Review*, pp. 964–985.

——— (2003): "Monetary policy evaluation with noisy information," *Journal of Monetary Economics*, 50(3), 605–631.

Orphanides, A., and J. Williams (2007): "Robust monetary policy with imperfect knowledge," *Journal of Monetary Economics*, 54(5), 1406–1435.

Piazzesi, M. (2005): "Bond Yields and the Federal Reserve," *Journal of Political Economy*, 113(2), 311–344.

Piazzesi, M., and E. Swanson (2008): "Futures prices as risk-adjusted forecasts of monetary policy," *Journal of Monetary Economics*, 55(4), 677–691.

Romer, C., and D. Romer (2004): "A new measure of monetary shocks: Derivation and implications," *American Economic Review*, pp. 1055–1084.

Stock, J. H., and F. Trebbi (2003): "Who Invented Instrumental Variable Regression?," *Journal of Economic Perspectives*, 17(3), 177–194.

Taylor, J. B. (1993): "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.

——— (2008): "The Costs and Benefits of Deviating from the Systematic Component of Monetary Policy," Keynote Address at the Federal Reserve Bank of San Francisco Conference on Monetary Policy and Asset Markets.

Tetlock, P. C. (2007): "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *The Journal of Finance*, 62(3), 1139–1168.

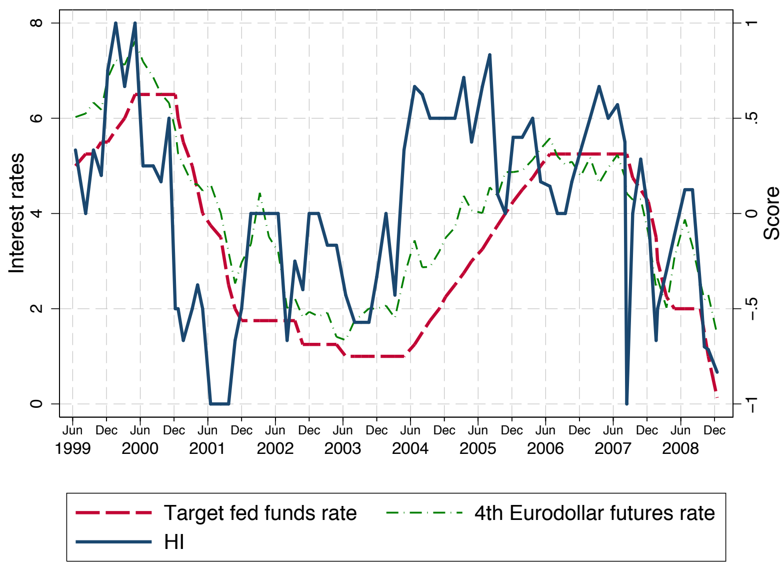
Turney, P. D. (2001): "Mining the Web for synonyms: PMI-IR versus LSA on TOEFL," *Proceedings of the Twelfth European Conference on Machine Learning*, pp. 491–502.

——— (2002): "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews," *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 417–424.

Turney, P. D., and M. L. Littman (2002): "Unsupervised Learning of Semantic Orientation from a Hundred-Billion-Word Corpus," National Research Council, Institute for Information Technology, Technical Report ERB-1094.

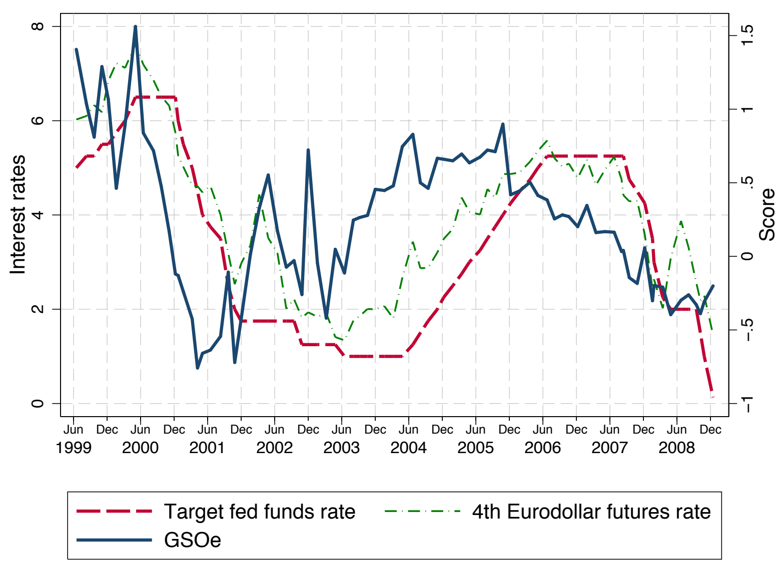
Woodford, M. (2005): "Central Bank Communication and Policy Effectiveness," Discussion paper, NBER Working Paper No. 11898.

Figure 1: Heuristic index (HI) and interest rates



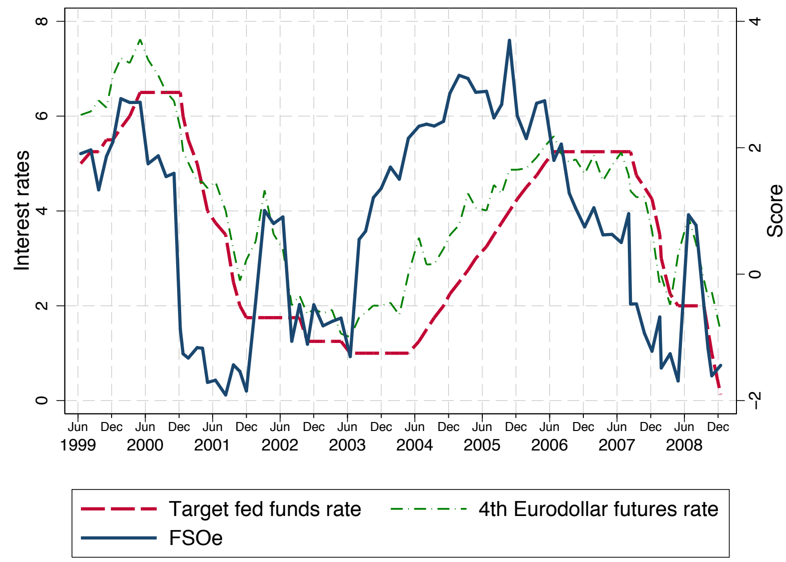
Notes: The lines in the chart show data on days of FOMC announcements after the release of the statement. The Heuristic index (HI) is defined in Equation (1). Information and definitions about the interest rate instruments is provided in Section 3.

Figure 2: Google semantic orientation score (GSO^e) and interest rates



Notes: The lines in the chart show data on days of FOMC announcements after the release of the statement. The Google Semantic Orientation score GSO^e based on the extended set of antonymies is defined in Section 2.1. Information and definitions about the interest rate instruments is provided in Section 3

Figure 3: Factiva semantic orientation score (FSO^e) and interest rates



Notes: The lines in the chart show data on days of FOMC announcements after the release of the statement. The Factiva Semantic Orientation score FSO^e based on the extended set of antonymies is defined in Equation (6). Information and definitions about the interest rate instruments is provided in Section 3

Table 1: Correlation matrix of levels of interest rates and semantic scores.

Variable	FFT	ED4	GSO^e	GSO^h	FSO^e	FSO^h
FFT	1.00					
ED4	0.91	1.00				
GSO^e	0.18	0.40	1.00			
GSO^h	0.30	0.37	0.16	1.00		
FSO^e	0.14	0.37	0.86	0.22	1.00	
FSO^h	0.35	0.46	0.38	0.31	0.41	1.00

Notes: FFT denotes the target federal funds rate and ED4 the futures implied rate on the 4^{th} Eurodollar quarterly contract. The remaining variables are levels of the Google- and Factiva-based semantic orientation scores defined on the extended set, and on the "hawkish-dovish", antonymies (respectively, GSO^e, GSO^h, FSO^e, and FSO^h). Values are measured on days of FOMC announcements after the release of the statement. Total number of observations for interest rates, the GSO^e and FSO^e score is 84. The numbers of observations for the GSO^h and FSO^h score are 82 and 68, respectively.

Table 2: Summary measures for the variables included in the interest rate response high-frequency regressions.

Change in Treasury yields						
	Δ 3-month	Δ 6-month	Δ 2-year	Δ 5-year	Δ 10-year	Δ 30-year
Mean	−1.15	−1.43	−1.07	−0.08	0.02	0.34
Median	0.00	−0.25	0.00	0.00	−0.20	−0.20
StDev	4.75	5.01	6.53	5.81	4.52	3.81
Min	−23.30	−24.30	−23.30	−18.74	−14.24	−6.50
Max	9.00	8.00	21.55	22.88	16.16	14.00

Change in Eurodollar futures rates				
	Δ ED1	Δ ED4	Δ ED6	Δ ED8
Mean	−1.57	−2.21	−1.97	−1.50
Median	−0.45	−1.75	−1.50	−1.00
StDev	7.64	8.52	8.16	7.70
Min	−38.00	−30.50	−29.50	−29.50
Max	17.75	24.50	25.00	25.50

Monetary policy surprise and change in Semantic Scores					
	MP	ΔGSO^e	ΔGSO^h	ΔFSO^e	ΔFSO^h
Mean	−2.24	−0.01	−0.01	0.03	−0.14
Median	−0.15	−0.01	−0.01	0.00	−0.34
StDev	9.85	0.29	0.46	0.41	1.19
Min	−46.50	−0.77	−1.50	−1.01	−2.38
Max	13.50	0.98	1.39	1.19	3.90
NoObs	82	82	80	82	58

NOTES: Number of observations for Treasury yields and Eurodollar futures rates is: 82. For interest rates, numbers are expressed as basis point changes during narrow temporal windows around FOMC announcements. Treasury yields are for the on-the-run issues. ED1-8 are future-implied rates on the nearest-eight quarterly Eurodollar futures contract. MP is the monetary policy surprise calculated as a (rescaled) difference in the current-month federal funds futures contract. ΔGSO^e is the change in semantic orientation score defined on the "extended "set of antynomies on Google data; ΔGSO^h uses the "hawkish-dovish "antynomy on Google data (see Section 2.1 for more detail). The scores ΔFSO^e and $\Delta FSOh$ are defined accordingly but use data from Factiva (see equation 6).

Table 3: Correlation matrix for the explanatory variables included in the interest rate response high-frequency regressions.

Variable	MP	ΔGSO^h	ΔGSO^e	ΔFSO^h	ΔFSO^e
MP	1.00				
ΔGSO^h	−0.02	1.00			
ΔGSO^e	0.26	−0.07	1.00		
ΔFSO^h	0.00	0.30	−0.02	1.00	
ΔFSO^e	0.08	−0.07	−0.21	0.11	1.00

NOTES: For variable definitions and number of observations included refer to Table 2.

Table 4: Regression Results for Treasury Yields

Δ Score:	None	Δ HI	ΔGSO^h	ΔFSO^h	ΔGSO^e	ΔFSO^e
Dependent Variable: Δ3-month yield						
MP	0.33	0.32	0.15	0.19	0.16	0.25
	[0.11]***	[0.11]***	[0.11]	[0.13]	[0.11]	[0.12]**
Δ Score	0.00	−0.57	0.03	0.00	−0.33	0.68
	[0.00]	[0.53]	[0.34]	[0.48]	[0.42]	[0.54]
R^2	0.47	0.48	0.47	0.46	0.47	0.47
Dependent Variable: Δ6-month yield						
MP	0.34	0.34	0.23	0.24	0.25	0.28
	[0.11]***	[0.11]***	[0.14]*	[0.14]*	[0.14]*	[0.12]**
Δ Score	0.00	−0.37	0.76	0.44	−0.65	1.18
	[0.00]	[0.66]	[0.48]	[0.48]	[0.50]	[0.53]**
R^2	0.46	0.46	0.46	0.41	0.46	0.47
Dependent Variable: Δ2-year yield						
MP	0.27	0.27	0.17	0.17	0.21	0.22
	[0.11]**	[0.12]**	[0.11]	[0.13]	[0.12]*	[0.10]**
Δ Score	0.00	0.41	1.98	1.82	−1.05	2.28
	[0.00]	[0.70]	[0.78]**	[0.63]***	[0.94]	[0.88]***
R^2	0.16	0.16	0.24	0.24	0.17	0.21
Dependent Variable: Δ5-year yield						
MP	0.07	0.08	−0.00	−0.01	0.05	0.05
	[0.09]	[0.10]	[0.10]	[0.11]	[0.12]	[0.10]
Δ Score	0.00	0.48	2.12	1.78	−1.14	1.66
	[0.00]	[0.63]	[0.97]**	[0.75]**	[0.86]	[0.84]**
R^2	0.02	0.02	0.12	0.12	0.04	0.04
Dependent Variable: Δ10-year yield						
MP	−0.01	−0.01	−0.07	−0.09	−0.04	−0.03
	[0.06]	[0.07]	[0.07]	[0.07]	[0.09]	[0.07]
Δ Score	0.00	0.40	1.70	1.12	−0.92	0.98
	[0.00]	[0.52]	[0.84]**	[0.66]*	[0.69]	[0.64]
R^2	0.00	0.01	0.09	0.08	0.02	0.01
Dependent Variable: Δ30-year yield						
MP	−0.12	−0.11	−0.13	−0.16	−0.12	−0.13
	[0.06]**	[0.06]*	[0.04]***	[0.05]***	[0.05]**	[0.05]**
Δ Score	0.00	0.73	0.75	0.74	−0.23	0.46
	[0.00]	[0.46]	[0.59]	[0.57]	[0.48]	[0.56]
R^2	0.09	0.13	0.11	0.20	0.09	0.10
N. Obs.	82	82	80	58	82	82
p-val. 3m=2y	.	0.04	0.01	0.02	0.39	0.02
p-val. 2y=30y	.	0.62	0.01	0.01	0.26	0.00

NOTES: Treasury yield responses during narrow temporal windows around FOMC announcements. Dependent variables are changes in Treasury yields at different maturities (horizontal panels). The dependent variables are the monetary policy surprise (MP), the change in the linguistic scores (columns), and a constant (not reported). Changes in interest rates are expressed in basis points and the standard deviations of the linguistic scores are normalized to one. See the footnote to Table 2 for variable definitions. The p-val. 3m=2y (2y=30y) is the p-value for the null that the 3-month and 2-year (2-year and 30-year) yield responses to the linguistic scores are identical. See Section 4 for detail about the calculation of the standard errors. *** significant at 1%, ** significant at 5%, *significant at 10%.

Table 5: Regression Results for Eurodollar futures rates

Δ Score:	None	Δ HI	ΔGSO^h	ΔFSO^h	ΔGSO^e	ΔFSO^e
Dependent Variable: ΔED1 futures rate						
MP	0.59	0.59	0.54	0.54	0.58	0.54
	[0.12]***	[0.12]***	[0.15]***	[0.16]***	[0.17]***	[0.14]***
Δ Score	0.00	−0.50	2.11	1.55	−1.04	−0.28
	[0.00]	[0.64]	[0.97]**	[0.69]**	[0.99]	[1.14]
R^2	0.59	0.59	0.62	0.62	0.59	0.59
Dependent Variable: ΔED2 futures rate						
MP	0.52	0.52	0.45	0.46	0.50	0.48
	[0.11]***	[0.12]***	[0.15]***	[0.15]***	[0.17]***	[0.13]***
Δ Score	0.00	0.02	2.36	1.97	−0.85	0.64
	[0.00]	[0.69]	[1.04]**	[0.69]***	[1.01]	[1.05]
R^2	0.45	0.45	0.50	0.52	0.45	0.45
Dependent Variable: ΔED4 futures rate						
MP	0.38	0.38	0.26	0.26	0.31	0.33
	[0.13]***	[0.14]***	[0.15]*	[0.16]	[0.18]*	[0.14]**
Δ Score	0.00	0.84	2.79	2.04	−1.11	2.02
	[0.00]	[0.93]	[1.23]**	[0.85]**	[1.22]	[0.88]**
R^2	0.19	0.20	0.27	0.25	0.19	0.21
Dependent Variable: ΔED6 futures rate						
MP	0.31	0.31	0.19	0.20	0.25	0.26
	[0.12]**	[0.13]**	[0.13]	[0.15]	[0.17]	[0.13]**
Δ Score	0.00	1.03	2.88	1.83	−1.31	1.82
	[0.00]	[0.86]	[1.28]**	[0.95]*	[1.24]	[0.97]*
R^2	0.14	0.15	0.24	0.20	0.14	0.15
Dependent Variable: ΔED8 futures rate						
MP	0.23	0.24	0.13	0.15	0.20	0.20
	[0.11]**	[0.13]*	[0.13]	[0.14]	[0.16]	[0.12]
Δ Score	0.00	1.19	2.74	1.59	−1.58	1.52
	[0.00]	[0.82]	[1.29]**	[0.95]*	[1.23]	[0.94]
R^2	0.09	0.11	0.18	0.14	0.10	0.10
N. Obs.	82	82	80	58	82	82
p-val. ED1=ED8	.	0.01	0.31	0.96	0.60	0.25

NOTES: Eurodollar futures implied rates responses during narrow temporal windows around FOMC announcements. Dependent variables are changes in Eurodollar implied rates at different maturities (horizontal panels). The dependent variables are the monetary policy surprise (MP), the change in the linguistic scores (columns), and a constant (not reported). Changes in interest rates are expressed in basis points and the standard deviations of the linguistic scores are normalized to one. See the footnote to Table 2 for variable definitions. The p-val. ED1=ED8 is the p-value for the null that the first and eighth Eurodollar rate responses to the linguistic scores are identical. See Section 4 for detail about the calculation of the standard errors. *** significant at 1%, ** significant at 5%, *significant at 10%.

Table 6: Federal funds rate forecasting regression

$$(\bar{r}_{t+\tau_{FF}} - r_{t^-}) = \beta_0 + \beta_1 \left(f_{t^-}^{\tau_{FF}} - r_{t^-}\right) + \beta_2 MP_t + \beta_3 \Delta FSO_t^e + \gamma X_{t^-} + \varepsilon_t^{\tau_{FF}}$$

FF:	FF2	FF3	FF4	FF5	FF6
Panel A Excluding other controls					
$\left(f_{t^-}^{\tau_{FF}} - r_{t^-}\right)$	1.07	1.13	1.19	1.23	1.28
	[0.10]***	[0.12]***	[0.12]***	[0.16]***	[0.19]***
MP	0.75	0.60	0.57	0.68	0.84
	[0.24]***	[0.39]	[0.45]	[0.48]	[0.51]
ΔFSO^e	4.72	11.55	14.52	19.30	23.32
	[1.99]**	[4.06]***	[5.03]***	[6.54]***	[7.83]***
R^2	0.86	0.80	0.79	0.77	0.74
Panel B Including term premium controls X_{t^-}					
$\left(f_{t^-}^{\tau_{FF}} - r_{t^-}\right)$	0.88	0.91	1.00	1.07	1.15
	[0.08]***	[0.11]***	[0.14]***	[0.20]***	[0.24]***
MP	0.68	0.42	0.34	0.41	0.58
	[0.20]***	[0.32]	[0.38]	[0.42]	[0.46]
ΔFSO^e	2.94	8.35	10.81	15.59	20.46
	[1.72]*	[3.00]***	[3.69]***	[5.83]***	[7.57]***
NBER rec. dummy	−14.65	−39.68	−44.40	−43.99	−36.52
	[7.11]**	[12.12]***	[15.19]***	[23.12]*	[32.93]
10y-2y slope	0.02	0.05	0.08	0.11	0.13
	[0.01]	[0.03]**	[0.04]**	[0.06]**	[0.07]*
BBB cred. spread	−5.44	−2.07	−1.76	−1.40	−1.53
	[3.68]	[5.34]	[6.20]	[7.38]	[9.14]
R^2	0.89	0.85	0.84	0.82	0.78
N. Obs.	82	82	82	82	82

NOTES: The columns in the Table report estimates of the model using quotes on the second (FF2) through the sixth (FF6) fed funds contract, which, given the schedule of FOMC meeting, correpond to forecasting horizons of about 1- to 5-months. The dependent variable $(\bar{r}_{t+\tau_{FF}} - r_{t^-})$ is the spread between the ex-post realized monthly average of the federal funds on each futures contracts' settlement month, $\bar{r}_{t+\tau_{FF}}$, and the intended fed funds rate ahead of the FOMC announcement t^-. The independent variables are the level of the futures-spot rate spread ahead the FOMC announcement, $\left(f_{t^-}^{\tau_{FF}} - r_{t^-}\right)$, the monetary policy surprise, MP_t, and the change in the Factiva semantic score, ΔFSO_t^e. Panel B includes three term premium controls measured ahead of the FOMC announcement: a dummy variable for NBER-dated recessions, the 10- to 2-year slope of the Treasury curve, and the 10-year BBB-rated corporate credit spread to Treasuries. All interest rates are expressed in basis points and the standard deviation of the linguistic scores is normalized to one. Newey-West standard errors with weighting and a truncation lag of 1-1/2 times each corresponding forecasting horizon reported in square brackets. For additional detail refer to Section 5.1. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 7: Eurodollar futures rate forecasting regression

$$(r_{t+\tau_{ED}} - r_{t^-}) = \beta_0 + \beta_1 \left(f_{t^-}^{\tau_{ED}} - r_{t^-}\right) + \beta_2 MP_t + \beta_3 \Delta FSO_t^e + \gamma X_{t^-} + \varepsilon_t^{\tau_{ED}}$$

ED:	ED1	ED2	ED3	ED4	ED5	ED6	ED7	ED8
Panel A Excluding other controls								
$\left(f_{t^-}^{\tau_{ED}} - r_{t^-}\right)$	1.09	1.29	1.34	1.25	1.29	1.40	1.50	1.67
	[0.19]***	[0.22]***	[0.25]***	[0.34]***	[0.35]***	[0.37]***	[0.38]***	[0.38]***
MP	1.20	0.85	0.42	1.95	2.68	3.11	4.82	4.21
	[0.41]***	[0.69]	[1.23]	[1.04]*	[1.24]**	[1.53]**	[2.43]*	[2.58]
ΔFSO^e	3.60	22.70	25.53	36.22	45.30	51.70	64.18	71.14
	[2.88]	[8.28]***	[8.29]***	[10.14]***	[11.39]***	[16.39]***	[20.08]***	[16.36]***
R^2	0.70	0.58	0.48	0.38	0.38	0.42	0.45	0.44
Panel B Including term premium controls X_{t^-}								
$\left(f_{t^-}^{\tau_{ED}} - r_{t^-}\right)$	0.94	1.06	0.68	0.37	0.29	0.37	0.28	−0.30
	[0.22]***	[0.25]***	[0.34]*	[0.31]	[0.41]	[0.41]	[0.49]	[1.12]
MP	1.04	0.66	0.27	0.70	0.79	−0.01	0.44	−1.68
	[0.46]**	[0.63]	[0.85]	[0.51]	[0.67]	[0.86]	[2.05]	[1.01]
ΔFSO^e	1.37	21.53	17.86	23.38	30.91	35.10	53.71	47.42
	[3.44]	[10.71]**	[7.06]**	[5.60]***	[8.14]***	[8.75]***	[9.08]***	[10.32]***
Recession Dummy	−17.59	−4.88	−5.82	18.97	−0.07	−20.53	−64.53	−101.51
	[15.03]	[44.13]	[61.76]	[59.08]	[77.56]	[86.89]	[69.48]	[37.27]***
10y-2y slope	0.04	0.13	0.36	0.67	0.87	1.07	1.36	2.06
	[0.04]	[0.10]	[0.18]*	[0.24]***	[0.28]***	[0.25]***	[0.32]***	[0.81]**
BBB cred. spread	−4.83	−19.76	−89.15	−198.91	−235.34	−271.22	−264.06	−241.38
	[6.23]	[10.24]*	[22.89]***	[45.27]***	[55.10]***	[60.42]***	[60.73]***	[61.30]***
R^2	0.73	0.63	0.64	0.70	0.75	0.78	0.79	0.78
N. Obs.	82	82	81	78	76	74	72	69

NOTES: The columns in the Table report estimates of the model using quotes on the first (ED1) through the eighth (ED8) Eurodollar futures contract, which, given the schedule of FOMC meeting, correpond to forecasting horizons of about 1-1/2 to 11-1/2 months. The dependent variable $(r_{t+\tau_{ED}} - r_{t^-})$ is the spread between the ex-post realized Libor rate on each futures contract settlement day, $t + \tau_{ED}$, and the intended fed funds rate ahead of the FOMC announcement, t^-. The independent variables are the level of the futures-spot rate spread ahead the FOMC announcement, $\left(f_{t^-}^{\tau_{ED}} - r_{t^-}\right)$, the monetary policy surprise, MP_t, and the change in the Factiva semantic score, ΔFSO_t^e. Panel B includes three term premium controls measured ahead of the FOMC announcement: a dummy variable for NBER-dated recessions, the 10- to 2-year slope of the Treasury curve, and the 10-year BBB-rated corporate credit spread to Treasuries. All interest rates are expressed in basis points and the standard deviation of the linguistic scores is normalized to one. Newey-West standard errors with weighting and a truncation lag of 1-1/2 times each corresponding forecasting horizon reported in square brackets. For additional detail refer to Section 5.1. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 8: Federal funds rate forecasting regression using after-announcement futures rates $f_{t^+}^{\tau_{FF}}$

$$(\bar{r}_{t+\tau_{FF}} - r_{t^-}) = \beta_0 + \beta_1 \left(f_{t^+}^{\tau_{FF}} - r_{t^-}\right) + \beta_2 MP_t + \beta_3 \Delta FSO_t^e + \gamma X_{t^-} + \varepsilon_t^{\tau_{FF}}$$

FF:	FF2	FF3	FF4	FF5	FF6
Panel A Excluding other controls					
$\left(f_{t^+}^{\tau_{FF}} - r_{t^-}\right)$	1.07	1.12	1.18	1.23	1.29
	[0.08]***	[0.10]***	[0.10]***	[0.13]***	[0.17]***
MP	−0.09	−0.12	−0.08	0.04	0.15
	[0.28]	[0.43]	[0.49]	[0.54]	[0.52]
ΔFSO^e	4.79	11.50	14.22	18.72	21.89
	[2.27]**	[4.21]***	[5.11]***	[6.60]***	[7.59]***
R^2	0.85	0.80	0.79	0.77	0.75
Panel B Including term premium controls X_{t^-}					
$\left(f_{t^+}^{\tau_{FF}} - r_{t^-}\right)$	0.88	0.90	0.99	1.07	1.17
	[0.08]***	[0.11]***	[0.13]***	[0.19]***	[0.23]***
MP	0.01	−0.15	−0.19	−0.13	−0.03
	[0.21]	[0.35]	[0.42]	[0.48]	[0.46]
ΔFSO^e	3.11	8.48	10.75	15.39	19.38
	[2.02]	[3.18]***	[3.79]***	[5.90]**	[7.25]***
NBER rec. dummy	−14.88	−38.96	−44.13	−42.47	−35.24
	[8.11]*	[13.75]***	[17.04]**	[24.20]*	[32.73]
10y-2y slope	0.02	0.05	0.08	0.11	0.13
	[0.02]	[0.03]**	[0.04]**	[0.06]**	[0.07]*
BBB cred. spread	−4.90	−1.71	−0.91	−0.60	−0.55
	[4.25]	[5.88]	[6.87]	[7.93]	[9.38]
R^2	0.88	0.85	0.84	0.81	0.79
N. Obs.	82	82	82	82	82

Notes: This Table reports parameter estimates for the model of Table 6 when controlling for level of futures rate in the futures-spot spread after, rather than ahead, the FOMC announcement, $\left(f_{t^+}^{\tau_{FF}} - r_{t^-}\right)$. For additional detail about the model specification refer to the footnote to Table 6. *** significant at 1%, ** significant at 5%, * significant at 10%.

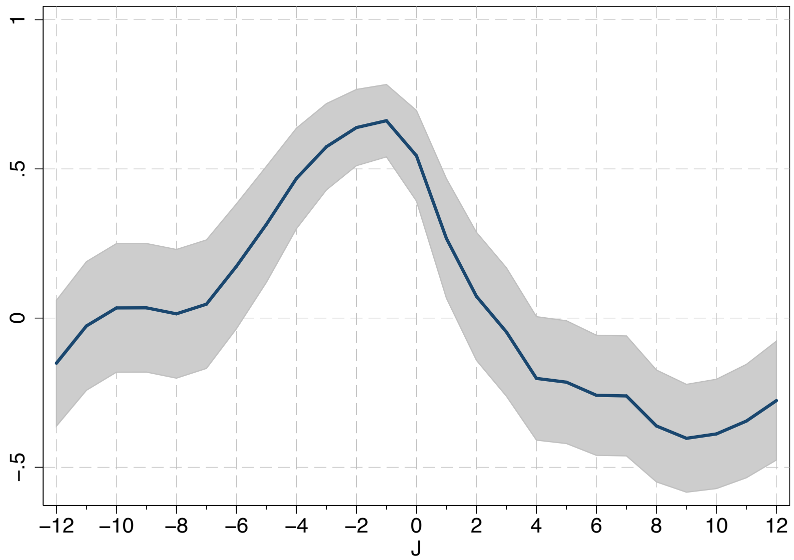
Table 9: Eurodollar futures rate forecasting regression

$$(r_{t+\tau_{ED}} - r_{t^-}) = \beta_0 + \beta_1 \left(f_{t^+}^{\tau_{ED}} - r_{t^-}\right) + \beta_2 MP_t + \beta_3 \Delta FSO_t^e + \gamma X_{t^-} + \varepsilon_t^{\tau_{ED}}$$

ED:	ED1	ED2	ED3	ED4	ED5	ED6	ED7	ED8
Panel A Excluding other controls								
$\left(f_{t^+}^{\tau_{ED}} - r_{t^-}\right)$	1.09	1.34	1.41	1.29	1.32	1.42	1.52	1.70
	[0.21]***	[0.21]***	[0.24]***	[0.33]***	[0.34]***	[0.36]***	[0.37]***	[0.36]***
MP	0.57	0.01	−0.32	1.46	2.28	2.83	4.43	3.92
	[0.52]	[0.65]	[1.23]	[1.03]	[1.24]*	[1.47]*	[2.26]*	[2.43]
ΔFSO^e	4.51	21.14	21.73	31.96	41.21	48.01	60.04	67.18
	[2.61]*	[9.00]**	[7.90]***	[9.97]***	[11.44]***	[16.07]***	[20.08]***	[16.38]***
R^2	0.69	0.58	0.50	0.39	0.39	0.42	0.45	0.44
Panel B Including term premium controls X_{t^-}								
$\left(f_{t^+}^{\tau_{ED}} - r_{t^-}\right)$	0.95	1.14	0.78	0.41	0.31	0.37	0.28	−0.32
	[0.27]***	[0.25]***	[0.34]**	[0.31]	[0.43]	[0.43]	[0.48]	[1.06]
MP	0.49	0.01	−0.16	0.54	0.70	−0.04	0.39	−1.66
	[0.57]	[0.59]	[0.94]	[0.53]	[0.59]	[0.81]	[1.95]	[1.09]
ΔFSO^e	2.18	20.79	15.54	21.98	29.91	34.54	53.35	48.07
	[3.15]	[10.89]*	[6.57]**	[5.31]***	[7.44]***	[8.35]***	[9.05]***	[9.16]***
NBER rec. dummy	−20.84	−5.58	−2.72	21.44	2.03	−19.01	−62.35	−103.60
	[14.91]	[40.22]	[59.78]	[59.07]	[79.10]	[87.58]	[72.28]	[30.52]***
10y-2y slope	0.04	0.12	0.33	0.65	0.86	1.07	1.36	2.08
	[0.04]	[0.10]	[0.18]*	[0.24]***	[0.29]***	[0.26]***	[0.32]***	[0.77]***
BBB cred. spread	−2.10	−15.03	−85.98	−196.97	−234.18	−270.64	−263.42	−241.79
	[8.51]	[11.37]	[22.58]***	[45.67]***	[55.69]***	[60.59]***	[60.31]***	[59.86]***
R^2	0.72	0.63	0.64	0.70	0.75	0.78	0.79	0.78
N. Obs.	82	82	81	78	76	74	72	69

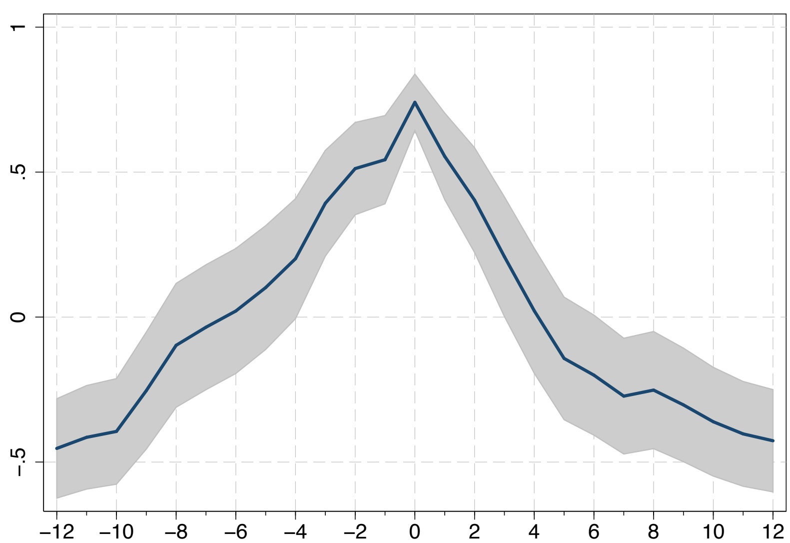
Notes: This Table reports parameter estimates for the model of Table 7 when controlling for level of the futures rate in the futures-spot spread after, rather than ahead, the FOMC announcement, $\left(f_{t^+}^{\tau_{ED}} - r_{t^-}\right)$. For additional detail about the model specification refer to the footnote to Table 7. *** significant at 1%, ** significant at 5%, * significant at 10%.

Figure 4: Cross correlations of the Taylor-rule residual and $FSO^e_{t\pm J}$



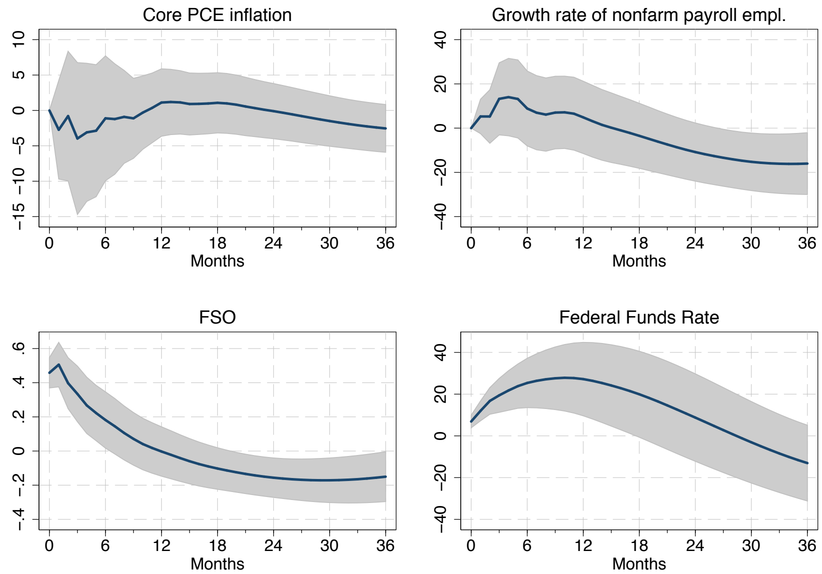
NOTES: The solid line represents the correlation coefficient between the Taylor-rule residual, defined in Section 5.2, and the Factiva Semantic Orientation score at different leads/lags $FSO^e_{t\pm J}$, with the unit interval defined as a quarter. Shaded areas represent two asymptotic standard-error confidence bands around the estimated correlations at each lead/lag.

Figure 5: Cross correlations of the Taylor-rule rate and $FSO^e_{t\pm J}$



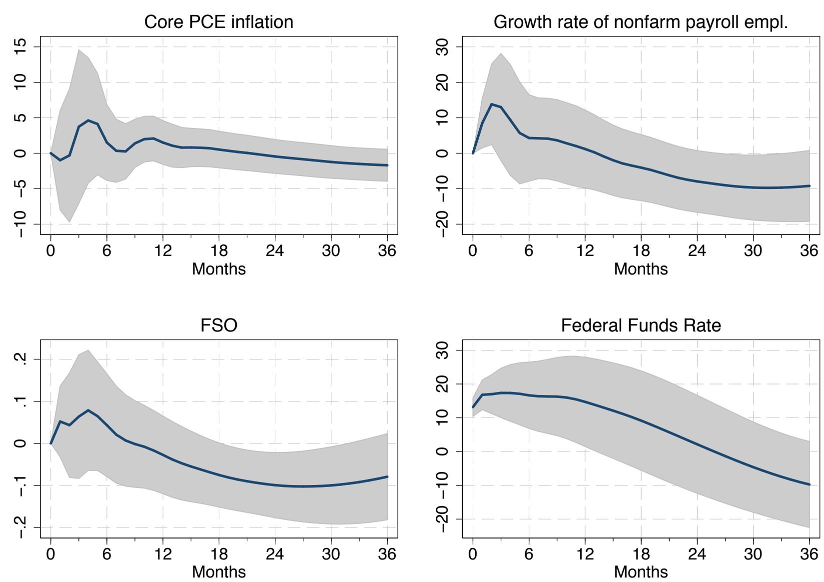
NOTES: The solid line represents the correlation coefficient between the predicted Taylor-rule rate, defined in Section 5.2, and the Factiva Semantic Orientation score at different leads/lags $FSO^e_{t\pm J}$, with the unit interval defined as a quarter. Shaded areas represent two asymptotic standard-error confidence bands around the estimated correlations at each lead/lag.

Figure 6: Impulse responses of the "core" variables to a FSO^e score shock in the VAR model



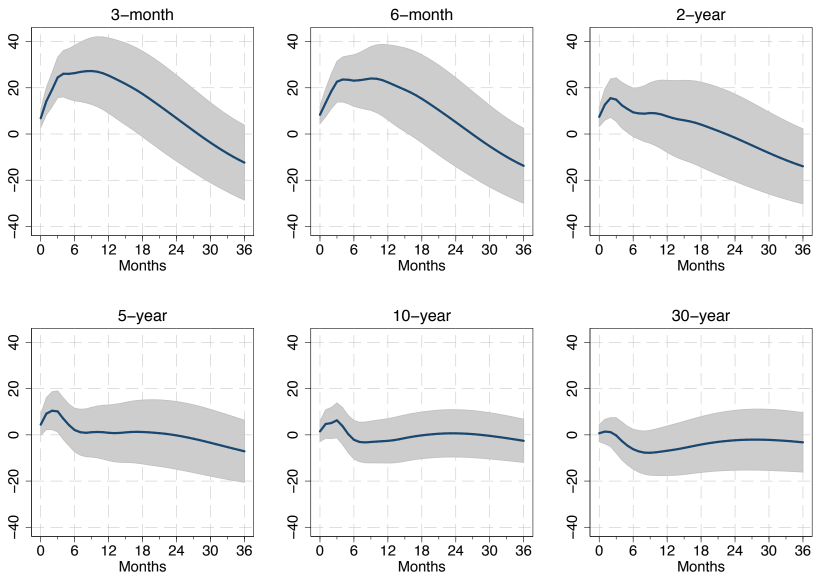
NOTES: The shock is identified in a recursive VAR with the "core" variables ordered as: inflation, employment, FSO^e score, federal funds rate. The VAR also includes Treasury yields, which do not identify the shock because of parameter restrictions. See Section 5.3 for additional detail. Shaded areas denote two-standard error bootstrapped confidence bands.

Figure 7: Impulse responses of the "core" variables to a federal funds rate shock in the VAR model



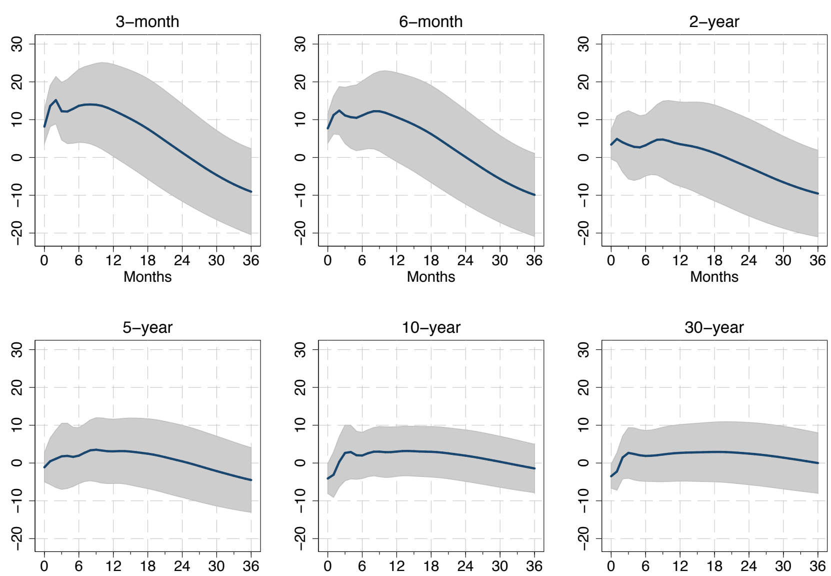
NOTES: The shock is identified in a recursive VAR with the "core" variables ordered as: inflation, employment, FSO^e score, federal funds rate. The VAR also includes Treasury yields, which do not identify the shock because of parameter restrictions. See Section 5.3 for additional detail. Shaded areas denote two-standard error bootstrapped confidence bands.

Figure 8: Impulse Responses of Treasury yields to an FSO^e score shock in the VAR models



NOTES: The shock is identified in a recursive VAR with the "core" variables ordered as: inflation, employment, FSO^e score, federal funds rate. Each impulse response corresponds to a different VAR model which, in addition to the "core variables", includes a different yield ordered last. Parameter restrictions are set such that the yield does not identify the shock and the shock is the same across specifications. See Section 5.3 for additional detail. Shaded areas denote two-standard error bootstrapped confidence bands.

Figure 9: Impulse responses of Treasury yields to a federal funds rate shock in the VAR models



NOTES: The shock is identified in a recursive VAR with the "core" variables ordered as: inflation, employment, FSO^e score, federal funds rate. Each impulse response corresponds to a different VAR model which, in addition to the "core" variables, includes a different yield ordered last. Parameter restrictions are set such that the yield does not identify the shock, and the shock is the same across specifications. See Section 5.3 for additional detail. Shaded areas denote two-standard error bootstrapped confidence bands.

Table 10: Forecast error variance decomposition for the "core" variables in the VAR model

Variable	Inflation		Employment		FSO^e		FFR	
Shock	FSO^e	FFR	FSO^e	FFR	FSO^e	FFR	FSO^e	FFR
Months Ahead								
3	0.2	0.0	0.8	3.6	90.0	0.7	34.9	55.2
	[1.7]	[1.6]	[1.9]	[3.4]	[7.7]	[2.1]	[10.0]	[9.8]
6	1.0	1.3	5.0	4.6	74.9	1.7	44.8	38.9
	[3.9]	[3.5]	[7.4]	[5.6]	[12.3]	[4.2]	[13.3]	[12.0]
12	1.1	1.5	4.4	3.1	54.4	1.3	45.1	23.1
	[4.4]	[3.5]	[8.7]	[4.9]	[14.6]	[3.8]	[16.1]	[11.6]
24	1.1	1.5	4.3	3.0	45.6	3.6	28.6	11.1
	[4.8]	[3.3]	[8.6]	[4.3]	[14.3]	[4.5]	[17.0]	[9.2]
36	1.6	1.7	11.3	5.7	47.3	6.8	21.9	8.7
	[4.9]	[3.2]	[9.6]	[5.1]	[14.5]	[5.5]	[16.1]	[8.4]

NOTES: For detail about the model specification refer to the footnote to Figure 6. The columns report the variance of the core variables explained by the shock to the federal funds rate (FFR) and the Factiva semantic orientation score (FSO^e). Numbers are expressed as percentage points. Bootstrapped standard errors reported in brackets.

Table 11: Forecast error variance decomposition for Treasury yields in the VAR models

Maturity	3-month		6-month		2-year		5-year		10-year		30-year	
Shock	FSO^e	FFR	FSO^e	FFR	FSO^e	FFR	FSO^e	FFR	FSO^e	FFR	FSO^e	FFR
Months Ahead												
3	37.9	29.6	37.6	21.5	20.6	2.3	10.0	0.1	3.2	1.7	0.3	1.6
	[10.1]	[9.4]	[10.1]	[8.6]	[9.6]	[4.2]	[7.9]	[2.4]	[5.2]	[3.1]	[3.0]	[2.6]
6	58.5	21.5	54.0	16.7	23.7	1.9	10.6	0.3	4.0	1.8	1.3	1.5
	[12.7]	[9.9]	[12.5]	[9.1]	[11.9]	[5.5]	[8.6]	[3.9]	[6.0]	[4.1]	[4.3]	[3.6]
12	54.0	16.3	44.6	12.2	13.2	1.7	4.7	0.8	2.9	1.7	6.1	1.1
	[16.0]	[10.3]	[15.2]	[8.8]	[11.4]	[5.6]	[7.0]	[4.8]	[5.4]	[5.0]	[7.1]	[4.4]
24	32.8	8.8	25.3	6.1	5.6	0.8	2.0	0.7	1.6	1.7	4.8	1.2
	[17.1]	[8.4]	[15.7]	[6.6]	[12.0]	[4.9]	[8.8]	[5.0]	[7.7]	[5.9]	[8.3]	[5.6]
36	25.7	7.2	20.2	5.5	5.8	1.7	2.0	0.7	1.3	1.4	3.7	1.0
	[16.3]	[7.7]	[15.1]	[6.2]	[11.7]	[5.0]	[9.3]	[4.6]	[8.8]	[5.6]	[9.2]	[5.7]

NOTES: For detail about the model specification refer to the footnote to Figure 8. The columns report the variance of the yields explained by the shock to the federal funds rate (FFR) and the Factiva semantic orientation score (FSO^e). Numbers are expressed as percentage points. Bootstrapped standard errors reported in brackets.