Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements*

David O. Lucca
Federal Reserve Bank of New York †

Francesco Trebbi
University of British Columbia and NBER‡

March, 2011

Abstract

We present a new method to measure central bank communication and apply it to statements by the Federal Open Market Committee. The measures are intuitive and capture significant information about future rate decisions by the Federal Reserve. We find that longer-dated Treasury yields mainly react to changes in communication around announcements. In lower frequency data, changes in the statements lead policy rate decisions by more than a year, and are a significant determinant of longer-dated Treasury yields. The statements contain information regarding both the predicted and the residual component of a Taylor rule model, and lead the residual component.

JEL Classification codes: E43, E52, E58.

*The authors would like to thank Matilde Bombardini, John Cochrane, Erik Hurst, Anil Kashyap, Monika Piazzesi, Jonathan Wright and Egon Zakrajšek 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 expressed in the paper are those of the authors and are not necessarily reflective of views at the Federal Reserve Bank of New York or 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.

†email: david.lucca@ny.frb.org.

‡email: ftrebbi@interchange.ubc.ca.

1 Introduction

This paper provides a new methodology to quantify the content of communication and presents its application to Federal Reserve announcements. The goal is not just to propose a transparent tool for the quantitative interpretation of verbal or written information, but also to provide a particularly stark example of the importance of communication in setting public policies. In the context of statements released by the Federal Open Market Committee (FOMC), we show how these measures extract information from the statements, explain financial markets response to their content, and characterize the information in the statements. With several different identification strategies we provide evidence that announcements matter substantially more than the immediate setting of short-term interest rates.

Our choice of application to central bank communication is driven by its increasing economic relevance. After years of intentional opacity, central banks around the world increasingly rely on communication with market participants to achieve their policy objectives (Mishkin [2007]). Since at least the early 1990s, central banks have signalled future policy rate actions, for example, by detailing their own projections for the economic outlook or for the targeted policy rate. Forward policy signalling allows central banks to affect long-term rates beyond the more conventional targeting of short-term interest rates (Bernanke [2004], Woodford [2005]). Indeed longer-term yields reflect, up to term premia, market expectations for the path of short-term rates, and thus react to credible communication of future policy decisions. With policy rates close to the zero-nominal lower bound in many advanced economies in the aftermath of the 2008 financial turmoil, communication has recently become an even more important policy instrument.[1] In contrast with such key role in the monetary policy toolkit, the literature on central bank communication is still in a relatively infant stage owing in part to the challenge of measuring verbal information directly in ways that are transparent, objective and replicable across researchers (see Blinder, Ehrmann, Fratzscher, Haan, and Jansen [2008] for a literature review). Statements released by the FOMC after its policy meetings are the primary mean of communication by the FOMC to market participants, and form an almost-ideal set of observations given the stability in the text structure

[1]After lowering their target rate close to the zero-lower bound with the onset of the turmoil several central banks have explicitly signalled their intention to maintain policy accommodation going forward. In the United States, the Federal Open Market Committee announced its intention to maintain "low levels of the federal funds rate for some time" in its December 18, 2008 policy statement, and for "extended period" in the March 18, 2009 statement. 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."

across observations, the number of available observations and their relative even spacedness in time.

We apply a set of intuitive, but information-theoretic based, tools from the field of computational linguistics designed to quantify language in terms of intensity and direction of meaning, or semantic orientation, and implement two classes of semantic scores. The first score—the Google semantic orientation score—is directly based on the text of 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.

To understand the Google semantic score, consider 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 one with which x and "dovish" occur. If x co-occurs more often with the word "hawkish" than with "dovish", then the sentence is intuitively more hawkish (and vice versa). Contributions in the linguistic literature provide an information-theoretic foundation to this intuitive approach (Church and Hanks [1990]). We implement the scores through searches on Google, since it is not possible to directly compute frequencies in the "population" of Internet webpages. Hit counts on joint searches (for example, a search of x and "hawkish") are empirical estimates of population frequencies and form the basis of the scores.

Because of the limited access of texts and the lack of control in Google's proprietary hit-count algorithms, we supplement the Google-based score with another implementation based on direct access of news from Factiva. In the Factiva-based analysis we first subset news from the database involving FOMC announcements around meetings. We then automatically analyze the resulting text and construct semantic orientation measures by comparing the frequency of different antonyms in the sentence database—e.g. the mutual association of word pairs such as "hawkish/dovish", "loose/tight"—to words indicating policy announcements such as "Fed" and "FOMC".

The automated approach that we present in this paper is novel to the economic literature and has several advantages relative to earlier work. First, it is a fully automated method replicable across researchers. Furthermore, by specifying an ex-ante metric on which we analyze meaning we depart from black-box methodologies, such as latent or content analysis methods, which produce results that are hard to interpret economically. Finally, because our algorithms are customized to the semantic analysis of monetary policy communication, they are better suited at capturing specialized meaning vis-a-vis alternative off-the-shelf tools that are calibrated on general language use.

After building these two automated scores, we study their properties as measures of monetary policy communication in a high- and low-frequency identification analysis. Using high frequency data around the release of FOMC announcements, we find that yields on longer-dated Treasuries mainly react to changes in the content of the statements (rather than unexpected changes in the fed funds target rate), 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 a univariate and a vector autoregression (VAR) model. The univariate model uses the semantic scores to forecast short-term rates at different horizons while conditioning on the information available to investors both right-ahead and after the policy announcements as implied by financial futures quotes. 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 the VAR estimates, a one standard deviation increase in the degree of "hawkishness" of the scores implies a hump-shaped response of the federal funds rate with a peak of about 30 basis point after about one year. The impulse responses also show that the semantic scores are a more important determinant of longer-dated Treasury yields than immediate policy actions. The results of the high- and low-frequency analysis support the view that the FOMC alters the content of the statements months ahead of taking policy rate actions, and consistently longer-term nominal Treasury yields respond to changes in the information in the statements.

Finally, in order to analyze the type of information contained in the statements, we decompose realized policy rates with a forward-looking Taylor rule. Based on this decomposition, we find that the semantic scores contain significant information regarding both the predicted and the residual component of the rule-implied interest rate decisions. More importantly, the semantic scores lead realized deviations from the Taylor-rule by a few quarters.

The work by Gürkaynak, Sack, and Swanson [2005] is an important earlier attempt to measure the content FOMC statements. Their measurement approach relies on latent factors of selected interest futures rates around policy announcements as communication proxies. Such an immediate approach is, however, limited by its indirectness. Econometric models measuring the impact of communication on interest rates based on these measures, in essence explain rates with rates. More importantly, as discussed in more detail in the paper, their latent factors, which in effect correspond to the "level" and "slope" of the yield curve, are known in the fixed income literature to explain nearly *all* rates variations at *all* times, irrespective of whether a policy announcements occur or not (Litterman and Scheinkman [1991]). An interpretation of these factors as measures of communication is thus not very straightforward. Other important contributions such as Romer and Romer [2004] or Bernanke, Reinhart, and Sack [2004] measure the content of verbal communication directly but rely on subjective ratings of text by the researchers.

A few interesting computational linguistics applications have appeared in the economic literature in recent years. Aside from the different empirical exercises, this paper departs from "black-box" off-the-shelf methods and uses a measurement approach that is at the same transparent and apt at capturing meaning of the short and semantically subtle FOMC statements. Boukus and Rosenberg [2006] apply latent semantic analysis (LSA) to FOMC minutes, an approach better suited to longer pieces of text, rather the concise structure of the FOMC statements. LSA identifies so-called "latent themes" from a set of texts and then classify them in terms of these themes. The economic interpretation of these latent factors is, however, not transparent. More intuitive linguistics indices, such as word counts, have been occasionally the focus of research in monetary economics, for example in Gorodnichenko and Shapiro [2007]. Other interesting applications of computational linguistics in the economic literature include Stock and Trebbi [2003], Antweiler and Frank [2004], Tetlock [2007], and Gentzkow and Shapiro [2006].

The remainder of this paper is organized as follows. In Section 2 we present the methodological description of the automated measures, and in Section 3 we discuss the data. In Section 4 we investigate Treasury yield responses to our scores, and in Section 5 we study low-frequency properties of our linguistic scores. Section 6 concludes.

2 Automated measures of the FOMC statement

Central banks around the world increasingly depend on communication to market participants to achieve their policy goals. Pareto-superior equilibria can be achieved when influencing private sector expectations by committing to specific policy rate paths (Woodford [2005]), and communication can help anchor private sector expectations around the central bank's long-run policy objectives (Bernanke [2004]).

In practice, because long-term rates depend on the entire expected path of short-term rates up to term premia, by signalling future policy rate intentions, monetary policy can more easily influence these rates, which are a key determinant of private sector allocations. 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 credible) is among the very few policy instruments available to exert monetary stimulus at the zero-nominal bound.[2]

In this section we build automated measures of the content of FOMC statements, which are the primary mode used by the Committee to communicate to investors. The statements are short in length: In our sample, the core of the statements, excluding the preamble describing the target rate decision and the concluding vote roll call, is composed on average of about six sentences, each of which about 25 words long. The statements express succinctly the FOMC's rationale for the most recent policy action, an assessment of the risks to its goals of "price stability and maximum sustainable employment" going forward, and at times explicit references of where the Committee expects its target for the federal funds rate to be in the near term. Examples of such more direct references to future policy rate decisions include statements that "[...]policy accommodation can be maintained for a considerable period", (August 12, 2003), that "the Committee believes that it can be patient in removing its policy accommodation" (January 28, 2004), and that "The Committee [...] continues to anticipate that economic conditions [...] are likely to warrant exceptionally low levels of the federal funds rate for an extended period." (December 16, 2009).

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 extract information about future policy rate actions. We define measures of the policy stance, or policy "hawkishness", and its intensity based on the FOMC statements, such that higher scores correspond to *hawkish* statements, pointing to increased chances of target rate increases, and lower scores to *dovish* statements, pointing to increased chances of target rate cuts.

The inherent difficulty of measuring words' meaning, discourse orientation, and intensity is the primary challenge in constructing these measures. For the sake of concreteness, consider the two phrases: "Pressures on inflation have picked up", (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'. 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.

As a benchmark consider the following scheme, which we will call heuristic index, or

[2] 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 empirical evidence against these stylized predictions.

score:

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

According to the operator defined in (1), the score $HI(X)$ would clearly be a 1, whereas $HI(X')$ would possibly be a 0. Such a heuristic approach, which is similar in spirit the approach used by Romer and Romer [2004] and Bernanke, Reinhart, and Sack [2004], has advantages and shortcomings. It is an intuitive and a simple measure of the orientation of a sentence. However, it coarsely approximates intensity and relies on subjective judgment by the researcher, limiting the interpretability and replicability across scorers.

Figure 1 shows the heuristic score (applying the scheme (1) to the sentences in each FOMC statement, and then averaging within each statement) alongside the intended (target) federal funds rate, and the rate implied by the fourth quarterly Eurodollar futures contract after the announcement, which is a market implied measure of short-term rates one-year ahead.[3] 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. The score is correlated with the contemporaneous level of the Eurodollar futures rate (about 50 percent in levels), although it appears to be more volatile than this rate at times.

We propose an alternative objective and automated method to similarly capture the semantic orientation the statement along a "hawkish–dovish" metric (or alternative metrics deemed appropriate—we experiment with six in total). Although a relatively new problem in economics, such measurement problems is commonplace in computational linguistics and statistical natural language processing (see Manning and Schütze [1999] for a review). Our approach is designed to fit the specific structure of FOMC statements (short text and substantial semantic finesse of the message).

We follow two different implementations to validate the analysis. 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 information retrieval on the Internet by Turney [2002]. The directness of the approach makes it easy to implement and transparent to interpret as compared

[3] The heuristic score reported is the consensus on the analysis of each statements by 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).

to other methodologies proposed in the computational linguistics literature.[4] 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 each FOMC announcement), drastically increasing the precision in measurement. In the remaining of this section we provide an overview of the methodology, while an extensive exposition is available in the online appendix.[5]

2.1 The Google semantic orientation score

We start by defining information theory's central concept of pointwise mutual information (PMI, see Manning and Schütze [1999]), which measures the association between concepts in a large corpus of reference text, in our case a string of text x and a key word, say the word "hawkish". If x is commonly interpreted as hawkish in the corpus, then x and the word "hawkish" should appear with a joint frequency, $\Pr(x \,\&\, \text{hawkish})$, greater than if the two were statistically independent concepts, in which case the joint would be the product of the marginals $\Pr(x)\Pr(\text{hawkish})$. PMI is simply the log-ratio of the joint to the marginals:

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

Following Turney [2002] and after applying PMI to x and the word "dovish", we define the semantic orientation score, measuring x's relative hawkishness, as $SO(x) = PMI(x, \text{hawkish}) - PMI(x, \text{dovish})$.

Following We implement the SO on the Internet, which is the largest corpus of text available. Because it is unfeasible to directly compute frequencies in the population of webpages, we implement the information retrieval process through hit counts on the Google search engine. The feasible estimator of the SO on Google using the "hawkish-dovish" word pair–denoted with superscript h—is then:

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

[4]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. See Turney [2001] and Turney and Littman [2002] for a comparison across the different approaches.

[5]Available at the authors webpages.

where hits(q) assigns 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. As an example, direct application of (3) to the strings X and X' through Google searches implies the higher score .98 for the hawkish sentence X, and the lower score -.53 for the relatively more dovish X'.

The antonymy "hawkish-dovish" (or others that we consider below) is arguably used in different contexts than monetary policy generating noise in the estimation of ξ. We abstract from this level effect by considering the measure:

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

in what follows, thus avoiding to interpret the level of the score. This does not limit our analysis that mainly relies on using the score in first differences. We implement the scheme (4) on each sentence of an FOMC statement by using automated lexical chunkers (as detailed in the online appendix) to obtain our search units x, which are in essence sub-sentences of text corresponding to the phrases in the sentence. We avoid using single words as our search units to maintain the original semantic content of the text, for example, by avoiding to separate adjectives from nouns or adverbs from verbs. The score for the statement is then the average $GSO^h(x)$ over all x's in the statement.

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

$$\Delta GSO^h_t = GSO^h_t - GSO^h_{t-1}. \tag{5}$$

This approximation will be true only when the score evolves according to a random walk, otherwise it implies measurement error in the construction of the shocks.

Finally we extend the SO-PMI measure to consider multiple sets of antonyms, not just the "hawkish"-"dovish" one. We implement an alternative score using six pairs of words for positive rate changes $\mathbf{P} = \{$hawkish, tighten, hike, raise, increase, boost$\}$, and negative ones $\mathbf{N} = \{$dovish, ease, cut, lower, decrease, loose$\}$. We define the score GSO^e, as the semantic orientation score based on the extended set of $\mathbf{P}$ and $\mathbf{N}$ antonymies. Taking each set of words as synonyms along our metric, the GSO^e is defined as the (log) ratio of the sum of the numerator and denominator terms (in the logs) in (4) across all words in $\mathbf{P}$ and $\mathbf{N}$ respectively.

The expanded list of antonyms increases the number of hits underlying the scores, but by using word pairs pertinent to both future and current policy actions as well as other contexts, more noise is added to the measure, which is particularly challenging when the measure is used in first differences.[6] Nonetheless, the resulting time series of score GSO^e appears reasonable in levels: It leads the policy rate by about two quarters (Figure 2) as accurately as, if not more than, the heuristic index, and displays a correlation with the fourth Eurodollar futures implied rate of about 40 percent (Table 1).

2.2 The Factiva semantic orientation score

Because we can only access the text of the corpus of webpages indirectly through Google searches, we cannot control the texts that are included in the searches, the time periods of reference of the texts, or the relevance of the matches obtained from the search. We circumvent these limitations by using 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.

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 articles on a 3-days window around the FOMC meeting starting on the day before, and ending on the day after, the announcement. The set of underlying sentences forms the corpus of text, $\mathbf{T}$, and includes 1,302,977 sentences, or about 15,500 per statement in our sample (82 statements in total).

Although this corpus of text is 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. 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$\}$. Define as $I[s, \mathbf{W}, \mathbf{W}']$ a function indicating whether the sentence s contains at least a word from list $\mathbf{W}$ and at least a word from list $\mathbf{W}'$. Given the set of sentences around the release date t, $\mathbf{T}_t$, the Factiva semantic orientation score for statement t is:

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

where $\mathbf{P}$ indicates the set of words associated with positive target rates movement (and

[6] 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.

N for negative ones). Equation (6) is the log-ratio of the frequencies of relevant sentences suggesting policy rate increases relative to those indicating rate declines. 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. In a similar fashion to the Google-based GSO^h score, we also consider a Factiva "hawkish-dovish" index, FSO^h.

Let us further indicate with $\mathbf{T}_{t^-}$ the subset of sentences that precede the FOMC announcement in the 3-day window, $\mathbf{T}_{t^+}$ the set of sentences that follow. The unexpected change in the information of the FOMC announcement at t is defined as difference in the score preceding and following the release:

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

This degree of precision can only be achieved in the Factiva implementation by using the date-time stamps of the articles in the database. 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. Having full control over the text search, alleviates this problem (in addition to controlling for the immediate policy action in the empirical specifications). 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. In addition, to improve on the precision of the measure we exclude 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"), as well as including in each set strict equivalent words (such as plurals, see the online appendix for additional detail). These refinements in the Factiva score improve the precision of the estimators, especially vis-a-vis the Google-based scores, but do not affect our empirical results in a fundamental way.

As reported in Figure 3, the resulting Factiva automated score FSO^e leads 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).

3 Data

The sample includes 82 FOMC statements starting in May 1999 and ending in December 2008.[7]

In the next Section, we study Treasury rate responses on narrow temporal windows around FOMC announcements starting 10 minutes before and ending 20 minutes after. The dependent variables are basis point yield changes for benchmark, or on-the-run Treasuries, with maturities ranging from 3-months to 10-years. In addition to changes in our semantic scores, the independent variables include the *unexpected* component of policy rate decisions—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.[8]

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. These changes are relatively small on average pointing to the fact that much of the policy announcements are already incorporated in prices. The lower-panel of Table 2 present summary measures for the dependent variables. The FSO^h score is only defined for about two-thirds of the observations included in the sample because of limited coverage. With the exception of the GSO^e, changes in the semantic scores are uncorrelated with monetary policy surprises highlighting how these measures are likely capturing information unrelated to current policy actions. The GSO^e score, instead, has a correlation of about 25 percent with the monetary policy surprises. As discussed in Section 2, this score is likely measured with error, as its level is influenced and partially reflects contemporaneous rate decisions taken at the meeting.

The univariate interest rate forecasting regressions in Section 5.1 include as controls, the monetary policy surprise as defined above, the 10/2-year slope of the Treasury yield curve 10 minutes ahead of the FOMC announcements, a credit spread between 10-year BBB-rated

[7] The starting point marks the date in which the FOMC begun systematically releasing statements after all meetings. For each full calendar year in our sample, the FOMC released 8 statements following scheduled meetings. 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 the Fed's lender of last resort function. As we focus on communication about future monetary policy rates, we only include statements for unscheduled policy meetings that discuss current or future policy rate decisions. In addition, due to missing financial quotes after the September 11 terrorist attacks, the September 17, 2001 statement is not included in our regressions.

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

corporate bonds on Treasuries at the close of the previous business day, and a dummy for NBER-dated recessions. The regressions also include futures rates as of 10 minutes before (and 20 minutes after) the announcement implied by the first 8 quarterly Eurodollar and the first 6 federal funds futures. Eurodollar contracts are settled on realized Libor rates at settlement dates, while federal funds rate futures settle on funds rate monthly average at expiry. The dependent variables in the regressions are chosen to match the settlement rates of the futures contracts included in the regressions.

The Taylor rule specification in Section 5.2 includes as controls 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 (as in Orphanides [2001, 2003]). These data are released to the public with a 5-year lag. We supplement the last three years 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) (data available at the Federal Reserve Bank of Philadelphia). The federal funds rate and semantic scores are average quarterly levels.

Finally the VAR analysis in Section 5.3 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 deflator. In addition the model includes monthly averages of par-yields on constant maturity Treasury yields from Gurkaynak, Sack, and Wright [2007].

4 High-frequency interest rate response regressions

In this section, we study nominal Treasury yield responses to changes in the information of FOMC statements about future policy rate decisions as measured by the semantic scores presented in Section 2. Following Gürkaynak, Sack, and Swanson [2005] and Fleming and Piazzesi [2005], we use a high-frequency identification approach to isolate the effects of the announcements from other same-day news or events, and regress yield changes from 10 minutes before to 20 minutes after the announcement on our semantic scores. For each Treasury maturity $i = 3m,\ 6m,\ 2y,\ 5y$ and $10y$ we regress yield changes Δy_t^i:

$$\Delta y_t^i = \beta_0^i + \beta_1^i \ MP_t + \beta_2^i \ \text{Score}_t + \varepsilon_t^i, \tag{8}$$

on the change in our semantic scores, and the monetary policy surprise, MP_t, which measures the unexpected policy rate action taken (or lack thereof) at the meeting as implied by futures quotes (see Section 3).

Given the novelty of our communication measurement approach, we begin by considering all four automated semantic orientation scores either defined on the "hawkish-dovish" antonymy on Google and Factiva, ΔGSO^h and ΔFSO^h, or on the expanded set of antonymies—ΔGSO^e and ΔFSO^e. We also consider as a benchmark the change in the human-generated heuristic score, ΔHI. For an easier interpretation of the results we standardize changes in all semantic scores in the regressions (coefficients thus measure basis point changes in yields per unit standard deviation increase in the scores).

As discussed in Section 2, changes in the Factiva scores are based on information collected in the 36 hours prior and following each announcement, and should thus proxy reasonably well for the unexpected information present in the statement.[9] Changes in the Google and heuristic scores are, instead, differences between current and lagged score values, and likely mix anticipated and unanticipated information. Under reasonable expectational assumptions, the resulting measurement error should downwardly bias the point estimates.

Regarding the error term in (8), differences in media and internet coverage over time likely lead to heteroschedasticity due to varying precision of the scores (mainly because statements in the earlier part of the sample received less and irregular coverage). We account for such heteroschedasticity by weighting observations through measurement precision proxies for each statement: the median number of hit searches for Google scores, and the total sentence number for Factiva scores.[10]

Table 3 reports the regression results. The horizontal panels in the Table show model estimates across yields, while the columns correspond to the different score measures included in each of the regression.

We start discussing two benchmarks. The first column reports estimates when only including the policy rate surprise MP_t (and a constant). The results are analogous to earlier findings (Gürkaynak, Sack, and Swanson [2005]) up to sampling differences. The unexpected policy rate action has significant effects on Treasury yields up to 2 years out—two to three basis point responses per unit standard deviation of MP_t that are statistically significant at conventional levels—and has no effect on longer dated Treasuries. The second column shows results for the human-generated score, HI. The HI score has no explanatory power whatsoever for Treasury yields in the sample. Some explanatory power of the score is lost starting in mid-2007 (not shown), likely owing to difficulties of the simple and rigid rule (1) at adapting to changes in the structure of more recent FOMC statements.

The remaining columns of Table 3 show results for the automated scores. Columns 3 and 4 report estimates for the Google- and Factiva-based scores (GSO^h and FSO^h) defined on

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

[10]Unweighted results present a loss in precision that occasionally reduces significance, but do not affect the qualitative results in high frequency. Results available from the authors upon request.

the "hawkish-dovish" antonymy. A unit standard deviation increase in ΔGSO^h leads to a hump-shaped yield response, with essentially no response for maturities up to six months, and about two basis point (and somewhat declining) responses beyond two year, which are statistically significant at conventional levels. The magnitudes of the responses are economically significant compared to the four to six basis point standards deviation in yields across maturities (Table 2). As discussed for the VAR model results below, following the initial shock, Treasury yields increase significantly and peak only after a few months after the initial shocks. The point estimates using the FSO^h are remarkably similar in magnitude (although with lower statistical significance) notwithstanding the much smaller number of observations included due to a complete lack of coverage in the Factiva corpus up to the end of 2003.

We now turn to parameter estimates for the scores defined on the expanded set of antonymies. The Google score on the enlarged set of antonymies, GSO^e, fails to predict yield changes across maturities (wrong-signed and insignificant coefficients). As previously discussed, the expanded list of antonyms enlarges the search hits, but also the measurement noise because of differences in the context of use of these words and the limited control on text on which the Google searches run. Results using the Factiva score on the expanded antonymies, FSO^e, are, instead, quantitatively very similar to those based on the "hawkish-dovish" antonymy, with a hump-shape response that peaks in the 2-year sector at slightly above two basis points. As previously discussed, we can directly calculate empirical frequencies on the texts underlying the Factiva corpus, and thus can better control for the pertinence of the text.

In sum, with the exclusion of the 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 nominal curve, policy communication affect longer-dated maturities. As discussed in the sections below, the information in the statements helps predict future policy rate actions. Under rational expectations, longer maturity yields reflect the anticipated future path of rate decisions, and thus are the most sensitive to the information in the statement.

Our results are also consistent with those of Gürkaynak, Sack, and Swanson [2005]. They proxy the surprise component of the announcement with a latent factor from futures rates, what they label as a "path surprise". Such an indirect approach is immediate but is limited by the fact that the communication measure is a combination of interest rates, which are of course affected by a number of things, not just communication. As an exercise, we reconstructed their "path surprise" using the same combination of futures rate changes (and factor rotations) on FOMC announcement days, but we then applied the same combinations on non-announcement days. On FOMC announcement days we replicate their results that the "path" factor explains nearly all of the variation in long-term rates. The same factor, which is their measure of communication, however, explains approximately the same variation in long-term yields on non-announcement days as well.[11] This result is not too surprising as, jointly with the monetary policy surprise, their measure, which is essentially a "slope" factor, span the first two principal components of the yield curve, known to capture yield variation at all times, with or without monetary policy announcements (Litterman and Scheinkman [1991]). Their indirect approach, although instructive, thus poses some challenges and requires associating their measure, or labelling it to FOMC communication, a step that our more direct measure does not require. In the working paper version of this paper, we also regressed eurodollar rates on our measures and found significant explanatory power futures that settle one- to one-half-years out, approximately the maturities from which Gürkaynak, Sack, and Swanson [2005]'s measure is derived. This explanatory power was, however, far from perfect. Notwithstanding these similarities around FOMC announcements, results in the next section suggest that futures rates do not fully incorporate the information about future rate decisions captured by our semantic scores.

5 Low-frequency results

In the remainder of the paper, we employ low-frequency data to study additional properties of our semantic measures in (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. For brevity we focus on the FSO^e score 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 its Google based counterpart, GSO^e, which, as previously noted, is likely affected by significant measurement error. The "hawkish-dovish" scores are estimated imprecisely until 2003 due to a very limited number of matches on the "hawkish-

[11] While Gürkaynak, Sack, and Swanson [2005] use intraday data on FOMC dates, we use daily data in this exercise matching the same January 1990 to December 2004 sample of the authors. We extract the factors from five daily interest rate changes: the second, third and fourth eurodollar implied futures rates, as well as changes in rates after the first and second upcoming FOMC meeting as implied by (rescaled) changes in federal futures rates corresponding to the months of the two meetings. The factors are constructed as detailed in the appendix of Gürkaynak, Sack, and Swanson [2005]. We first obtain two factors using the principal components method, and then orthogonally rotate the two factors so that one factor, the "path surprise", is orthogonal to the monetary policy surprise. We obtain the rotation matrix and principal components on FOMC dates and then apply these same combination to interest rate changes on all other days to obtain the factors. Although they use intraday changes, we replicate their results very closely on FOMC dates using factors obtained from daily changes. Detailed results are available from the authors.

dovish" antonymy, allowing an analysis on a very limited sample size.[12]

5.1 Forecasting short-term interest rates

In this section, we study the in-sample predictive power of the semantic orientation scores for future realized short-term rates. Specifically we assess if more hawkish FOMC statements—as measured by the scores—predict (in a Granger sense) higher short-term rates, as postulated in the previous Section. For each forecasting horizon τ_n, the forecasting model that we estimate is:

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

The model controls for all information available to investors ahead of the FOMC announcement as implied by futures quotes taken 15-minutes before the announcement, $f_{t^-}^{\tau_n}$ (below, we also consider quotes taken 15 minutes after, $f_{t^+}^{\tau_n}$) while the dependent variable is the rate, $\tilde{r}_{t+\tau_n}$, on which each futures contract settles. The forecasting horizons, τ_n, range between one and five months for federal funds futures (second to sixth contract). Because the liquidity of these futures declines sharply after the first few month expiries, we also consider the more liquid Eurodollar futures for horizons between 1-1/2 months (first Eurodollar contract) and 1 year and 11-1/2 months out (eighth contract).[13]

Because of interest rate persistence, we follow the literature testing the expectation hypothesis (for example, Fama and Bliss [1987] and Campbell and Shiller [1991]) and subtract spot interest rate levels, r_{t^-}, ahead of the meeting from the realized, $\tilde{r}_{t+\tau_n}$ and the futures rate $f_{t^-}^{\tau_n}$. The dependent variable in (9) is thus the difference between realized and current spot rates, whereas the independent variable is a futures-spot spread.

Under forecasting efficiency of futures 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

[12] 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. 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.

[13] 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. 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 the day of the FOMC meeting t. 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 target rate decision using the monetary policy surprise, MP_t, and the new information in the statement with the change in the FSO^e_t score. The coefficient β^n_3 measures the basis point response in realized short-term rates of a unit standard deviation unexpected increase in the FSO^e_t score.

Although we cannot reject the null of forecasting efficiency of futures rates when only including the futures-spot spread and a constant, previous literature finds excess returns on fed funds and Eurodollar futures to be strongly countercyclical in the 1990s, likely due to time-varying risk premia. Following Piazzesi and Swanson [2008] we control for a set of term premium proxies, X_{t^-}, in the regression: 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 for NBER-dated recessions.[14]

Because of overlapping forecasting horizons in (9) 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 in each calendar year, we compute Newey and West [1987] HAC standard errors with truncation lags equal to 1-1/2 times the forecasting horizon, τ_n. For consistency we also keep the weighting procedure described in Section 4 when estimating (9).[15]

The parameter estimates of the forecasting model for the federal funds and 3-month Libor rates when controlling for futures quotes ahead the announcement, $f^{\tau_n}_{t^-}$, are shown in the upper panels ("A") of Tables 4 and 5, respectively. The columns in the tables report estimates for different forecasting horizons corresponding to the maturity of the futures rate controls: the second to sixth fed funds futures rate, a forecast horizon of 1 to 5 months, in Table 4; the first to eighth Eurodollar futures rate, corresponding to 3-month realized Libor starting about 1-1/2 months to about 2 years out. The parameter estimates for a constant and the term premium controls are omitted for brevity in the tables. As shown in Table 4, the sensitivity of future rates to a one standard deviation unexpected shock to FSO^e build up monotonically over time and range between about 3 basis points 1-month out and 20 basis

[14]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 sample.

[15]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 (9). 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.

points after five months (significant at conventional levels beyond the third contract). Panel A of Table 5 report parameter estimates for forecasting regression that control for eurodollar futures rates. The estimated coefficients on the standardized ΔFSO^e, range between about 1 basis point at one month and 45 basis points almost 2-year out, with a peak coefficient of about 55 basis points about 1-2/3 years out (seventh Eurodollar contract). All parameter estimates beyond the first contract significant at conventional levels. In sum, an unexpected shock to the FSO^e score predicts an economically large response of future short-term rates, with a peak response after more than a year after the initial shock. We find consistent results below when using a VAR identification approach.[16]

As noted in Section 4, although the semantic scores contain significant predictive power for yield changes around FOMC announcements, when controlling for the monetary policy surprise, their explanatory power is far from perfect, suggesting a potentially different information content of the semantic scores and futures implied rates. As a final exercise we re-estimate the model (9) using quotes taken after, rather than before the FOMC announcements, to construct the futures-spot controls. The bottom panels ("B") of Tables 4 and 5, show parameter estimates when substituting $f_{t^+}^{\tau_n}$, where t^+ is 20 minutes after the FOMC announcement at date t, to $f_{t^-}^{\tau_n}$ in (9). Comparing estimates in each column of Panel B to those in Panel A, it is interesting to note that the coefficients on ΔFSO^e are little affected by the inclusion of futures quotes taken after the release of the FOMC statement. In theory one would expect futures quotes after the release of the statement to contain all relevant information to predict realized rates. However, based on this (in-sample) analysis, implied rates do not appear to fully achieve such task, perhaps due to the presence of term premia. Movement in premia around the announcement may garble the information content of these quotes in predicting future realized rates, and thus proxying the content of the statement with futures quotes may not fully capture all relevant information content in the statements.

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

[16] 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 (first fed fund and Eurodollar futures contracts). Based on the reported standard errors, we cannot the null of forecasting efficiency the futures-spot spread, $(f_{t^-}^{\tau_n} - r_{t^-})$ for forecasting horizons below 6-months, while efficiency is rejected for longer horizons. Regarding the term premium controls (point estimates shown in the working paper version of this paper), the NBER recession dummies enter with a negative (and significant) sign under 4-month horizons (5^{th} fed funds contract), coefficients on the slope of the yield curve are positive and significant across (nearly all) maturities, and the corporate bond spread enters negatively and significant beyond 5-months (3^{rd} Eurodollar contract).

looking Taylor rule model for the federal funds rate. We then 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, or "interest rate gap". Finally we compute cross-correlations between the two components and the semantic score at different leads and lags, and evaluate the information in the statements based on these correlations.

We consider a Taylor-rule specification that incorporates partial interest rate adjustment (Clarida, Galí, and Gertler [2000]), and 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 try 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 only available to the public with a 5-year lag. We supplement the Greenbook forecasts starting in 2005 with forecasts of inflation and real GDP from the Survey of Professional Forecasters (SPF), and of potential GDP from the Congressional Budget Office (CBO).[17]

To capture additional macroeconomic variation in the Taylor rule, we estimate parameters on the longer sample September 1987 to December 2008, a period that covers the tenures of chairmen Greenspan and Bernanke, and then use these estimates to decompose policy decisions in our shorter sample. The model specification closely follows Orphanides [2003]:

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

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

[17] Because these forecasts 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.

exclusion of the lagged interest rate term:

$$i_t^T \equiv \beta_0 + \beta_\pi \ \pi_{t+3}^a + \beta_{y^a} \ \Delta y_{t+3}^a + \beta_y \ y_{t-1}.$$

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:

$$\hat{i}_t^T = .79_{[.05]} \ i_{t-1} - .04_{[.13]} + .49_{[.10]} \ \pi_{t+3}^a + .43_{[.07]} \ \Delta y_{t+3}^a + .25_{[.04]} \ y_{t-1}.$$

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.[18] In addition, β_{y_a} 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 predicted values for ϵ_t and i_t^T, we calculate the cross-correlation functions $\text{Corr}(\hat{\epsilon}_t, FSO_{t-J}^e)$ and $\text{Corr}(\hat{i}_t^T, FSO_{t-J}^e)$, which are shown in Figures 4 and 5. As it is generally the case, the "Taylor rule rate" explains a very large portion of the policy rate variation ($\bar{R}^2$ of about 98 percent). Given the results in the previous sections, it is not too surprising that the the semantic score FSO^e is significantly correlated with $\hat{i}_t^T$ (with a contemporaneous peak of about 65 percent). More interestingly, however, the FSO^e score is as much as 75 percent correlated with the residual component, and leads the residual by one to two quarters. The interest rate gap is commonly interpreted as a shock to policy makers' preferences, for example, because other controls affect their policy decisions at times, or due to shifts in the weights on the policy objectives. Although we cannot disentangle between these or other interpretations, our results suggest that the FOMC communicates significant information regarding such deviations in the statements ahead of the actual policy rate actions.

Because longer-term yields (see Section 4) and other macroeconomic variables generally respond to such information, the econometricians' information set in conventional VAR models (for example, Christiano, Eichenbaum, and Evans [1999]) is possibly smaller than that of the agents making the shocks "non-fundamental" in such models. The next section considers a small-scale VAR that controls for the additional information through our scores.

[18] Point estimates on the inflation term guarantee stability in economies in which interest rates are set using these forwarding looking rules as 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 $\beta_\pi/(1-\alpha) = 2.6$.

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 in a VAR multivariate framework, where shocks are identified in a very different approach. In this model, we can also study interest rate responses for months, and not just on tight temporal windows as in Section 4, after the initial shocks.

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 policy shocks, while keeping the number of parameters low, we estimate five 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. 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 three months and ten 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 with a standard recursiveness assumption: First, the change in the macroeconomic variables, $\mathbf{X}_t$, respond with a lag 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. Although the main qualitative findings in this Section do not hinge on the ordering within $\mathbf{S}_t$, we find it to be preferable to the alternative as, following the discussion in Section 5.2, we interpret the FSO^e as partially indicating a change in the preferences of policy makers to which the funds rate should be allowed to respond immediately.

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.[19] The structural form of the VAR models can be written as:

[19] 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].

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

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

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

and the matrix σ is diagonal. The diagonal terms in the matrices $\mathbf{a}_{11}$ and $\mathbf{a}_{22}$ of (12) are equal to one, and the innovations $\boldsymbol{\epsilon}_t^{\mathbf{X}}$, $\boldsymbol{\epsilon}_t^{\mathbf{S}}$ and $\epsilon_t^{R^i}$ in (11) are structural shocks. It is important to note that, because of the zero elements in (12), the structural shocks of the policy block $\boldsymbol{\epsilon}_t^{\mathbf{S}}$ do not depend on the Treasury yield included in each VAR model, so that each VAR model considered identifies the same monetary policy shocks. Based on the AIC, 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.[20]

In the discussion below we focus on responses of the federal funds rate and Treasury yields, and refer to the working paper version for detail about all other responses. Figures 6 and 7 show responses of these variables to unexpected unit standard deviation shocks in their innovations, $\boldsymbol{\epsilon}_{\mathbf{t}}^{\mathbf{S}}$. Positive innovation to both the funds rate and the FSO^e score are interpretable as contractionary monetary policy shocks; whereas positive innovations to the federal funds rate directly feed into higher short term rates, a shock to FSO^e affects allocations as more "hawkish" statement is subsequently followed by a target rate increases. All responses shown in the charts are absolute basis point deviations from the unshocked values, and shaded areas are two-standard error bootstrapped confidence bands.

The federal funds rate displays a hump-shaped response to a shock in the FSO^e score with a peak of about 30 basis points about a year after the shock. After a shock to its own innovation, instead, the funds rate peaks only at about half as much and almost immediately. The forecast error variance decompositions, imply that the shock to the linguistic score accounts for a significant portion of the federal funds rate variance 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.

These results confirm those 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

[20] 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.

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.

We now turn to the responses of Treasury yields. As shown in Figure 6, yields at all maturities rise on impact after a positive innovation to the FSO^e score, with magnitudes that generally decline with longer maturities. The responses of the 3- 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. Ten-year Treasury responses are never statistically different from zero. The yield responses to a federal funds rate shock are shown in Figure 7. 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 longer maturities 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.

In sum we find that short-, medium-term Treasury yields and the federal funds rate increase after a VAR identified 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 roughly one-year after the shock. The responses of medium-term Treasuries are less long lived. With the exception of the immediate responses of short-dated Treasuries, these results indicate 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 the past decade.

6 Conclusions

This paper presented a novel approach to measure the content of central bank communication regarding future policy rate decisions, and applied these measures to FOMC statements, which are the primary mean of communication of the Committee to market participants.

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 limited attention in the empirical monetary economics literature. The importance of communication in determining the term structure of nominal rates, along with the leading properties of the semantic measures relative to rate decisions, suggests that econometrician may omit significant information available to economic agents when identifying monetary policy shocks in standard monetary models (for example, in canonical VARs as in Christiano, Eichenbaum, and Evans [1999]).

Work in the economic and finance literature (such as Gürkaynak, Sack, and Swanson [2005]), as well as central bankers and market practitioners, have often 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 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 at least in predicting future realized policy decisions. More fundamentally proxying the content of communication with a set, or combination, of interest rates, ultimately requires attributing movements in these yields to communication, a step that, as discussed in the paper, is at best hard to test.

Relative to heuristic methods, our approach has the advantage of being unsupervised, intuitive, and replicable across researchers. At the same time our method departs from other off-the-shelf computational linguistics methods in being customized to the discussion of monetary policy and on relying on large quantities of reference texts from the internet and news media. Both features are important to effectively capture meaning in the concise, and semantically subtle, texts of FOMC statements.

The information used to construct our semantic measures relies on commentaries by members of the press, market practitioners, and investors. In this sense, our measures can 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.

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.

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

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

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.

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.

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.

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

Litterman, R., and J. Scheinkman (1991): "Common factors affecting bond returns," *The Journal of Fixed Income*, 1(1), 54–61.

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.

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.

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.

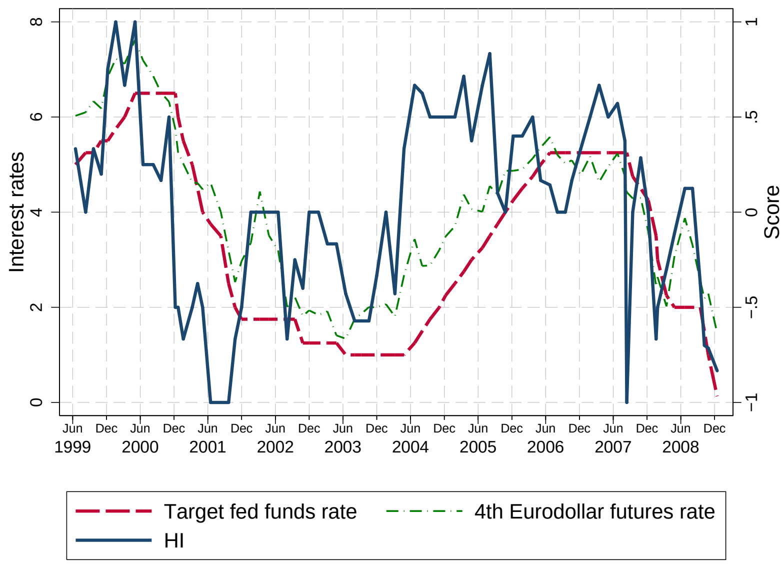
Turney, P. D. (2001): "Mining the Web for synonyms: PMI-IR versus LSA on TOEFL," *Proceedings of the Twelfth European Conference on Machine Learning*, pp. 491–502.

——— (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.

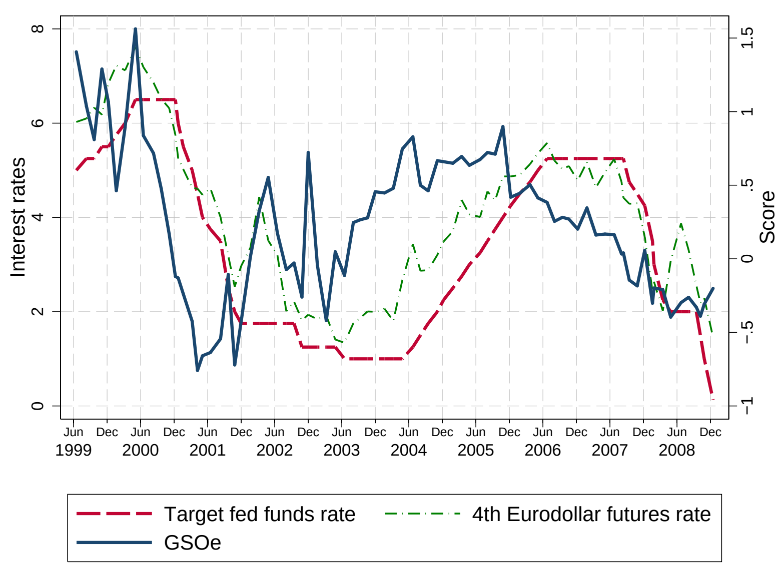
Woodford, M. (2005): "Central Bank Communication and Policy Effectiveness," Discussion paper, NBER Working Paper No. 11898.

Figure 1: Heuristic index (HI) and interest rates



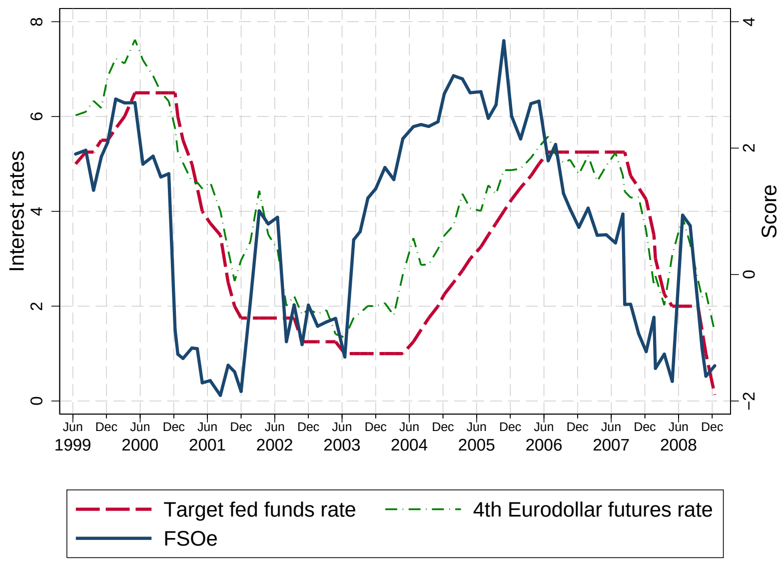
NOTES: The lines in the chart show data on days of FOMC announcements after the release of the statement. The Heuristic index (HI) is defined in Equation (1). Information and definitions about the interest rate instruments is provided in Section 3.

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
Mean	−1.15	−1.43	−1.07	−0.08	0.02
Median	0.00	−0.25	0.00	0.00	−0.20
StDev	4.75	5.01	6.53	5.81	4.52
Min	−23.30	−24.30	−23.30	−18.74	−14.24
Max	9.00	8.00	21.55	22.88	16.16
	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	82.00	82.00	80.00	82.00	58.00

NOTES: Number of observations for Treasury yields and rates is: 82. Interest rate entries are expressed as basis point changes in a 30 minute temporal window around FOMC announcements. Treasury yields are for the on-the-run issues. 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: 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
	[0.04]***	[0.04]***	[0.04]***	[0.04]***	[0.04]***	[0.04]***
Δ Score	.	−0.57	0.03	0.00	−0.33	0.68
	.	[0.39]	[0.41]	[0.52]	[0.53]	[0.45]
Dependent Variable: Δ6-month yield						
MP	0.34	0.34	0.23	0.24	0.25	0.28
	[0.04]***	[0.04]***	[0.05]***	[0.05]***	[0.05]***	[0.04]***
Δ Score	.	−0.37	0.76	0.44	−0.65	1.18
	.	[0.42]	[0.47]	[0.64]	[0.62]	[0.49]**
Dependent Variable: Δ2-year yield						
MP	0.27	0.27	0.17	0.17	0.21	0.22
	[0.07]***	[0.07]***	[0.06]***	[0.06]**	[0.07]***	[0.06]***
Δ Score	.	0.41	1.98	1.82	−1.05	2.28
	.	[0.67]	[0.64]***	[0.82]**	[0.87]	[0.68]***
Dependent Variable: Δ5-year yield						
MP	0.07	0.08	−0.00	−0.01	0.05	0.05
	[0.07]	[0.07]	[0.06]	[0.06]	[0.06]	[0.06]
Δ Score	.	0.48	2.12	1.78	−1.14	1.66
	.	[0.65]	[0.63]***	[0.79]**	[0.87]	[0.68]**
Dependent Variable: Δ10-year yield						
MP	−0.01	−0.01	−0.07	−0.09	−0.04	−0.03
	[0.05]	[0.05]	[0.05]	[0.05]*	[0.05]	[0.05]
Δ Score	.	0.40	1.70	1.12	−0.92	0.98
	.	[0.51]	[0.52]***	[0.63]*	[0.72]	[0.54]*
N. Obs.	82	82	80	58	82	82

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. See Section 4 for detail about the calculation of the standard errors. *** significant at 1%, ** significant at 5%, *significant at 10%.

Table 4: Federal funds futures forecasting regression

	FF2	FF3	FF4	FF5	FF6
Panel A Futures rate measured before meeting: $f_{t^-}^{\tau_{FF}}$					
$\left(f_{t^-}^{\tau_{FF}} - r_{t^-}\right)$	0.88	0.91	1.00	1.07	1.15
	[0.08]***	[0.11]***	[0.14]***	[0.20]***	[0.24]***
MP_t	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]***
Panel B Futures rate measured after meeting: $f_{t^+}^{\tau_{FF}}$					
$\left(f_{t^+}^{\tau_{FF}} - r_{t^-}\right)$	0.88	0.90	0.99	1.07	1.17
	[0.08]***	[0.11]***	[0.13]***	[0.19]***	[0.23]***
MP_t	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]***
N. Obs.	82	82	82	82	82

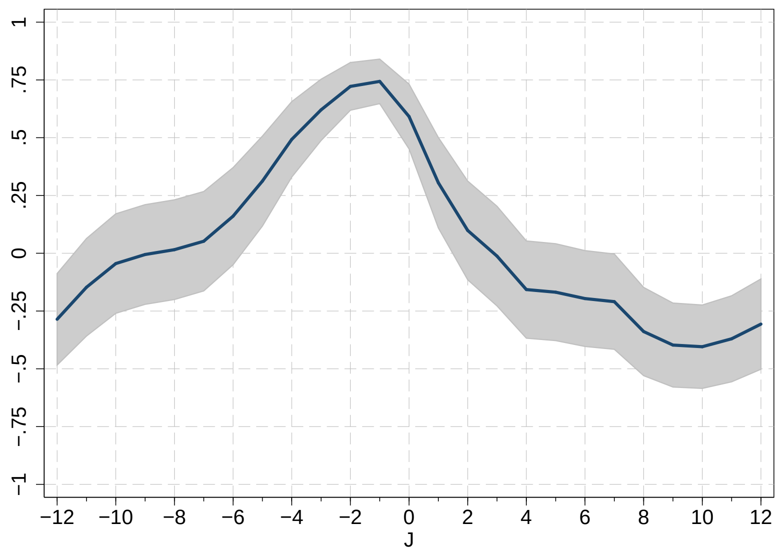
NOTES: The columns report parameter estimates in specifications that include quotes on the second (FF2) through the sixth (FF6) fed funds futures contract (forecast horizons of 1- to 5-months). The dependent variable $(\bar{r}_{t+\tau_{FF}} - r_{t^-})$ is the spread between the realized average funds rate on the settlement month, $\bar{r}_{t+\tau_{FF}}$, and the funds rate ahead of the FOMC announcement t^-. In Panel A, the futures-spot rate spread is measured ahead the FOMC announcement, $\left(f_{t^-}^{\tau_{FF}} - r_{t^-}\right)$, while in Panel B after the announcement, $\left(f_{t^+}^{\tau_{FF}} - r_{t^-}\right)$. The variable MP_t is the monetary policy surprise, while ΔFSO_t^e is the change in the Factiva semantic score. The regression includes a constant and term premium controls (estimates not shown):a dummy 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. Interest rates are 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 5: Eurodollar futures forecasting regression

	ED1	ED2	ED3	ED4	ED5	ED6	ED7	ED8
Panel A Futures rate measured before meeting: $f_{t^-}^{\tau_{ED}}$								
$\left(f_{t^-}^{\tau_{ED}} - r_{t^-}\right)$	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_t	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]***
Panel B Futures rate measured after meeting: $f_{t^+}^{\tau_{ED}}$								
$\left(f_{t^+}^{\tau_{ED}} - r_{t^-}\right)$	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_t	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]***
N. Obs.	82	82	81	78	76	74	72	69

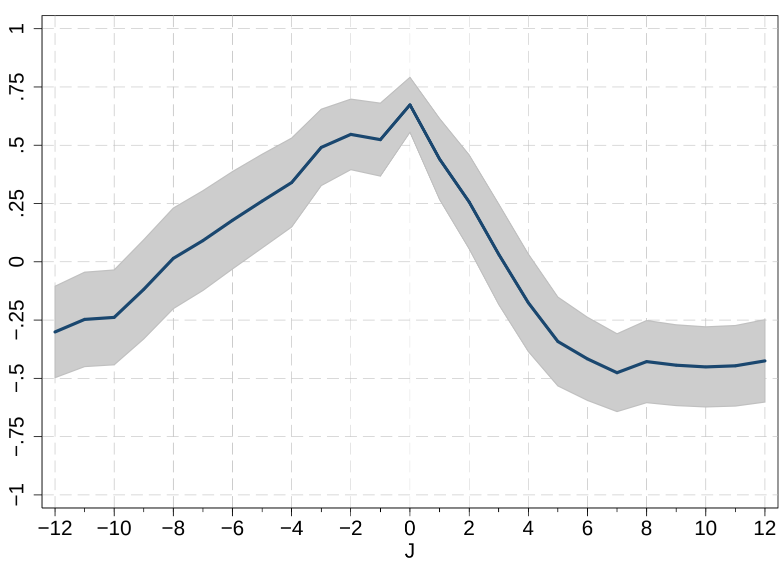
Notes: The columns report parameter estimates in specifications that include quotes on the first (ED1) through the eighth (ED8) Eurodollar futures contract (forecast horizons of 1-1/2 to 11-1/2 months). The dependent variable $(r_{t+\tau_{ED}} - r_{t^-})$ is the spread between the ex-post realized Libor rate on settlement days, $t + \tau_{ED}$, and the Libor rate ahead of the FOMC announcement, t^-. In Panel A, the futures-spot rate spread is measured ahead the FOMC announcement, $\left(f_{t^-}^{\tau_{ED}} - r_{t^-}\right)$, while in Panel B after the announcement, $\left(f_{t^+}^{\tau_{ED}} - r_{t^-}\right)$. The variable MP_t is the monetary policy surprise, while ΔFSO_t^e is the change in the Factiva semantic score. The regression includes a constant and term premium controls (parameter estimates not shown): a dummy 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. Interest rates are 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%.

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 in the VAR models of Treasury and fed funds rates to an FSO^e score shock

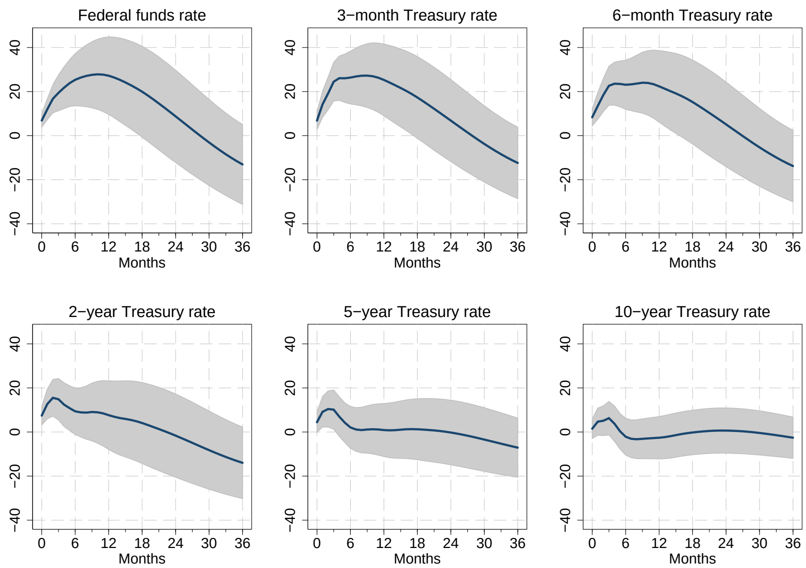
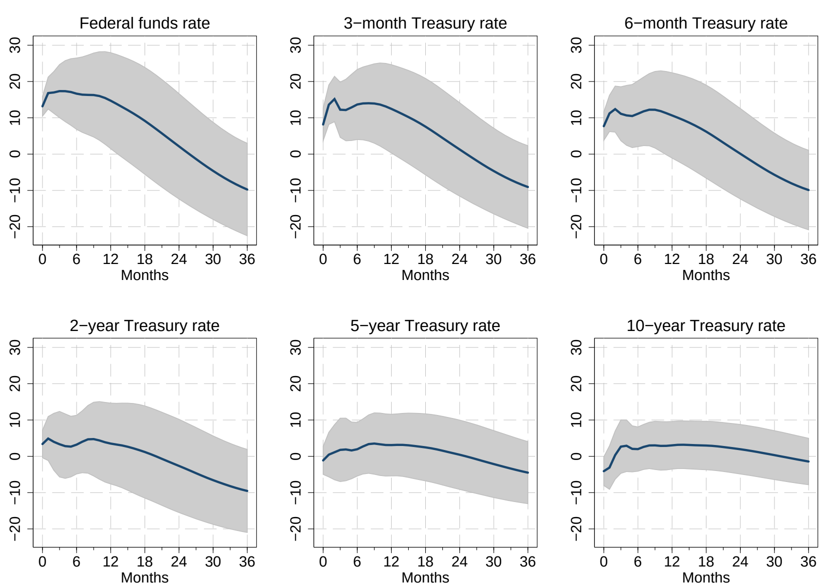


Figure 7: Impulse responses in the VAR models of Treasury and fed funds rates to a funds rate shock



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.