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

David O. Lucca
Federal Reserve Board †

Francesco Trebbi
University of Chicago and NBER‡

First Draft: August 2007
This version: **April 7, 2008**

Abstract

We present a new automated, objective and intuitive scoring method to measure the content of central bank communication about future policy rate moves. We apply the methodology to statements released by the Federal Open Market Committee (FOMC) after monetary policy meetings. Using intra-day financial data, we find that short-term nominal yields on Treasury securities respond to changes both in policy rates and the content of the statement. By contrast, medium- and long-term yields only react to changes in communication. Using lower frequency data, we find that changes in the content of the statements lead policy rate moves by about six months both in univariate and vector autoregression models. These results are consistent with the view that the FOMC releases information about future policy actions in the statement and market participants include this information when pricing short- and medium-term securities.

The paper discusses the interplay between central bank communication and policy rate moves.

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 Zakraj šek for their comments and suggestions. Isaac Laughlin provided outstanding research assistance. Hal Varian and Google University Research Program provided generous access to Google Search technology. The views and analysis set forth are solely those of the authors and do not indicate concurrence by other members of the Board of Governors of the Federal Reserve System. Francesco Trebbi gratefully acknowledges support by the Initiative on Global Markets and the Mutch Family Faculty Research Fund at the Graduate School of Business, the University of Chicago.

†Division of Monetary Affairs, email: david.o.lucca@frb.gov.

‡Graduate School of Business, email: ftrebbi@chicagogsb.edu.

1 Introduction

This paper presents a class of automated measures of monetary policy communication that contribute to the burgeoning literature of information transmission between central banks and other economic agents. The results of the empirical analysis highlight the important role played by monetary policy communication in affecting medium- and long-term risk-free nominal interest rates, and provide insights regarding the interaction between central banks' communication and their setting of short-term policy rates.

The paper applies the new automated measurement approach to the content of the statements released by the Federal Open Market Committee (FOMC) after monetary policy meetings. These statements are of particular interest for two separate reasons. First, the texts of these statements represent an almost-ideal set of observations for an empirical analysis of communication: The structure of the text is fairly comparable over time; the statements are available for a relatively long period of time; finally, the release dates are fairly equally spaced in time. Second, according to popular financial press and findings of previous literature, financial market participants pay very close attention to the content of these statements, and, at times, even small changes in the wording of the statements elicit significant reactions of asset yields and prices both in the U.S. and in foreign financial markets.[1]

After years of intentional opacity on the part of monetary policy authorities, the last fifteen years have witnessed extensive and increasing efforts on the part of central banks to improve transparency of their communication to market participants.[2] Such efforts have often been frustrated by the lack of quantitative assessment of the role of such communication. Although theoretical contributions to the analysis of information and communication have been substantial, empirical research has not produced comparable insights, mostly because of the complexity involved in analyzing quantitatively flows of verbal information in a manner that is at the same time objective, intuitive, and replicable.[3] Indeed, text and words are not intuitively susceptible of quantification in terms of intensity and direction of meaning, which in what follows will be referred to as semantic orientation. Different people tend to subjectively interpret non-quantitative information and the same set of words can have very different meaning and intensity depending on the context of use. In this paper, we borrow a set of tools from computer science and computational linguistic that are specifically designed to address these measurement issues and that are based on an intuitive, but information-theoretic based principle.[4] Given two opposing concepts (a dichotomy, say, "hawkish" vs. "dovish"), the semantic orientation of a sentence X (say, "Pressures on inflation have picked

[1]See for example Bernanke, Reinhart, and Sack [2004], or the recurring Wall Street Journal column "Parsing the Fed", which analyzes the content of each sentence of the statement relative to the most recent one "for clues about where interest rates may be headed."For foreign assets responses see, e.g., Hausman and Wongswan [2006].

[2]See, for example, the discussion in Bernanke [2004b] and Greider [1989].

[3]For theoretical contributions related to monetary policy see: Morris and Shin [2005], Woodford [2005], and Svensson [2006]. For communication within the firm see: Crémer, Garicano, and Prat [2006].

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

up") can be measured by the relative frequency with which X and the word "hawkish" jointly occur, and the frequency with which X and the word "dovish" jointly occur. If the string "Pressures on inflation have picked up" co-occurs more often with the word "hawkish" than with "dovish", then it seems intuitive to attribute to that sentence a relatively more hawkish score (and vice versa). It is through the use of Internet search engines that the empirical estimation of those joint frequencies can be easily implemented. It suffices to consider hits counts on joint searches (e.g., run a search of sentence X and the word "hawkish") that are then mapped into joint frequencies. As we show in the methodological section of the paper, simple search routines in Google can help assign quantitative scores to the content to FOMC statements. We show in the following sections how communication scores based on mutual associations on web pages capture well the policy stance of the central bank and lead movements in the policy rate—the federal funds rate—by several months.

This approach has several advantages relative to previous literature. First, it does not rely on subjective ratings of text of researchers, like for instance in Romer and Romer [2000] or Bernanke, Reinhart, and Sack [2004] , and is based on objective search routines. At the same time, by specifying an ex-ante metric along which we analyze the content of the statement—in particular, we focus on the degree of "hawkishness" of the statement as predictor of future policy rate hikes—we depart from black-box methodologies, such as factor analysis methods, which deliver findings that are hard to interpret economically, and are often silent about policy communication prescription.[5] The same type of comment also applies to methods based on word frequencies and counts, which are inherently ambiguous on the meaning or orientation of the statements. Finally, this approach makes the analysis easy to automatize and to replicate by other researchers.

After discussing and implementing the scoring technique, we study the empirical properties of the semantic scores using a high- and low-frequency identification analysis. We find that yields on short-term Treasury securities respond both to changes in policy rates and changes in the content of the FOMC statements over short time windows around the time of FOMC announcement. Instead, yields on medium- and long-term securities only respond to changes in the content of the statements, with 2-year Treasuries having the largest responses. Using data at monthly and intermeeting frequencies, we then analyze the relation between central bank communication—as measured by the automated scores—and policy rates using two models: a univariate model specification and a vector autoregression (VAR) model. The univariate model is used to directly forecast short-term rates—the federal funds and 3-month Libor rate—at different time horizons using the semantic scores while conditioning on all information available to market participants about the future rates right before the policy announcements (as implied by quotes on futures contracts delivering at the corresponding time horizon).[6] The VAR model, instead, includes the federal funds rate and the semantic scores, as well as measures of inflation and economic activity. Both models show that the

[5] See among others Gürkaynak, Sack, and Swanson [2005b], as well as Boukus and Rosenberg [2006] for an application of latent semantic analysis to FOMC minutes.

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

content of the statements has considerable forecasting power for future short-term rates. Parameter estimates of the univariate model show that the content of the FOMC statements has predictive power for short-term rates up to one year out. In addition, the estimates of the VAR model imply that a one standard deviation unexpected increase in the policy stance of the statements yields to higher policy rates, with a peak of about 20 basis point after about seven months. This evidence supports the view that the Federal Open Market Committee modifies the content of the statement several months ahead of taking policy rate actions. Consistently, medium- and long-term Treasury yields respond to changes in the content of the statements.

Finally, we assess the type of information contained in the statements, as measured by the semantic scores, under the assumption that the FOMC followed the policy prescriptions of a Taylor [1993] type rule in setting short-term policy rates during Chairman Greenspan's and Bernanke's tenures. After estimating the parameters of the rule, we find an economically large correlation between the semantic scores and both the systematic and unsystematic components of policy rates. In particular, the information contained in the FOMC statements seems to apply both to the current non-systematic component of policy decisions, that is the component of these decisions not directly explained by developments in inflation and economic activity, and to subsequent systematic components of policy rate moves.

The rest of paper is organized as follows. In Section 2 we discuss the role of central banks communication as an expectation management tool. In Section 3 we report the methodological description of the class of automated communication measures employed in the paper. In Section 4 we present the data. In Section 5 we investigate the effects of communication on asset yields using high-frequency data. In Section 6 we analyze the low-frequency properties of our linguistic scores: we first evaluate their forecasting power in an univariate setup; we then analyze the empirical link between the linguistic scores and the systematic and non-systematic components of future and current policy decisions, as determined by Taylor rules; we finally analyze the relation between the scores and policy actions in a recursive VAR model. Section 7 concludes.

2 Central bank communication

The statements released by the Federal Open Market Committee after policy meetings, are among the most important tools used by its members to communicate to financial market participants and other agents. In order to put the empirical analysis that follows in a clearer perspective, this Section briefly discusses the role of the statements, and of central bank communication more in general, as monetary policy instruments.

Following Bernanke [2004a], central banks use their communication to achieve several objectives. First, systematic communication increases the accountability of monetary authorities, a particularly important goal given their central role and political independence. Second, because unexpected policy actions come with large asset price reactions, as well as with large "winners" and "losers", communication may help improve the overall stability of the financial system. Third, and most importantly for this paper, central banks use communication as an expectation management tool. Indeed, monetary authorities use their communication to manipulate market participants' views about the likely path of future policy actions—and align these views to their own—and, more in general, communication is used to improve the understanding of long-run policy objectives, helping central banks to achieve their long-term goals, most notably anchoring inflation expectations.

As discussed in Bernanke [2004a] and Woodford [2005], central bank communication increases the effectiveness of monetary policy by influencing long-term interest rates beyond what can be achieved through movements in short-term policy rates. This simply follows from the fact that long-term rates depend, up to risk- and term-premia, from the expected future path of short-term rates. Because long-term rates often have a more important role in households' and businesses' investment decisions—e.g. through mortgage rates—central bank communication can enhance the effectiveness of monetary policy, by affecting market expectations.[7]

More in general expectation management is a particularly important medium- and long-term policy tool. Indeed, agents' expectations concerning future policy rate moves determine in part the path of prices and quantities in the economy, and if central banks can influence these expectations by committing to a specific policy path, they can achieve pareto-superior equilibria (Woodford [2005]). With time-varying objectives and preferences that are not directly observed by other agents, central banks can use communication to signal the future policy path, therefore aligning agents' expectations to the central bank's preferred policy path. Of course, under rational expectations, of future policy moves have to be confirmed in equilibrium, implying an empirical correlation between words and future actions.[8]

Consistent with an interpretation of central bank communication as an expectation management tool, in the next Section we construct a set of measures of the content of the FOMC statements that attempt to extract information from the statements about future policy rate actions.[9]

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

[8]In Section 6.1, we find that the content of the FOMC statement has indeed significant predictive power for future changes in the federal funds rate.

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

3 Measurement of communication: the FOMC Statement

This section sets forth the quantitative analysis of the FOMC statement. The core of a FOMC statement is composed of five sentences—5.35 on average for the period between May 18, 1999 and August 7, 2007—each sentence composed of twenty-five words—24.64 on average over the same period— expressing succinctly the FOMC's rationale for the, or lack of, policy action and an assessment of the risks to its goals of "price stability and maximum sustainable employment."[10]

We restrict our objective to defining a measure of the stance—and intensity of the stance—of FOMC statements regarding future policy rate hikes or cuts. In other words, we aim at measuring the degree of "hawkishness" or "dovishness" of the views expressed in the statements about the FOMC's own assessment of the risks to the goals of its dual mandate. There are two reasons for following this approach. First, the measure is directly suggested by the "balance of risk" of the statements and, second, generally speaking, market participants interpret the statements in a similar manner according to popular financial press.[11] The measure should produce a high score for a *hawkish* statement—i.e., forecast increases in the intended federal funds rate—and a low score for a *dovish* statement—i.e., forecast decreases of the federal funds rate.

The challenge is that words are inherently difficult to measure and so are their meaning, discourse orientation, and intensity. For the sake of concreteness, suppose we were set to analyze the difference between two sentences: "Pressures on inflation have picked up", statement of March 22, 2005—call this string of text X—and "Inflation pressures seem likely to moderate over time", December 12, 2006—call it string X'.[12] Read in their entirety, the former sentence can be interpreted as being hawkish and the latter as being dovish. While most would agree that X delivers a stronger indication than X' of an outlook of rising inflation and, possibly, future hikes in the policy rate, no clear metric exists prima facie to assess the two.

In order to emphasize the advantages of our methodology, we discuss next a simple approach to go about the problem: We assign to each sentence a subjective, heuristic score. For instance, consider the following scheme, which we will call heuristic score (HI), that applies to $x \in \{X, X'\}$:

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

[10]In the paper we will refer to the FOMC statement as corresponding to its "core" text, that is, excluding the preamble that describes the policy rate action, and the concluding description of the voting roll call. Our sample is restricted to the period between May 18, 1999 and August 7, 2007 because earlier statements were not systematically released by the FOMC after all policy meetings.

[11]In the period between May 1999 and February 2000 statements included a "policy bias" as opposed to a "balance of risk" assessment. For further discussion see Kohn and Sack [2003].

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

According to the operator in (1), the score $HI(X)$ would clearly be one, whereas $HI(X')$ would possibly be a zero. A heuristic approach as the one just described has advantages and shortcomings.[13] It is an intuitive and simple measure of the orientation, or hawkishness of a sentence. However, it coarsely approximates for intensity and relies on an arbitrary and subjective judgment of the researcher and, thus, it is difficult to interpret or to replicate across scorers. Consider, for instance, the alternative interpretation of the score of X' as being equal to -1 on the grounds that lower inflationary pressures tend to be associated with a fall in aggregate demand. In addition the categories in (1) might not always be mutually exclusive as, for example, in the case of a period of stagflation.

Yet one might ask: how well does such a score describe the monetary policy stance of the FOMC? By applying the scheme (1) to the set of sentences of each FOMC statement in our sample, and after averaging within each statement, we obtain the heuristic measure reported in Figure 1.[14] The Figure shows the intended (target) federal funds rate and the heuristic score. From a descriptive standpoint, the heuristic score appears to lead the policy rate by about one-or-two quarters, consistent with the idea that the score measures the stance of policy. Statements with a dovish stance appear to lead subsequent policy rate cuts, while statements with a hawkish stance have been followed by subsequent rate hikes.

The new approach that we propose in the rest of this Section is to assign to each sentence an objective, automated score. The difficulty here is generating an automated algorithm, able to capture the semantic stance of the statement or one of its parts along the hawkish–dovish metric. Although it is a relatively new problem in economics, the unsupervised and automatic measurement of the intensity or the semantic orientation of a text is commonplace and a long-standing issue in computational linguistics and natural language processing—scientific fields at the intersection between computer science and linguistics.[15] Here, we follow the approach initially proposed by Church and Hanks [1990] and generalized to the analysis of semantic orientation through information retrieval by Turney [2002], and Turney and Littman [2003]. The directness of the approach makes it relatively easy to implement and clear to understand compared to other methodologies.[16]

[13]See Romer and Romer [2000] and Bernanke, Reinhart, and Sack [2004] for applications of subjective measures in the analysis of the Fed's policy stance and communication using FOMC minutes and statements, respectively.

[14]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 be little or no ambiguity of interpretation.

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

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

Assume that the metric that we wish to define can be properly characterized along a simple dichotomy, say scoring of a string of text x—either a word or a sentence—on an hawkish–dovish scale: That is, we wish to create a score defined over the real line whose values depend on the hawkishness of the given string of text x.[17] We begin by defining a measure of association between concepts. If the meaning of a string x can be commonly interpreted as hawkish, then x and the word "hawkish" should show a degree of positive statistical dependence in a sufficiently large corpus of text. In other words, the string x and the word "hawkish" should appear in a language with a joint frequency, $\Pr(x \,\&\, \text{hawkish})$, larger than if the two strings were statistically independent concepts in which case the joint frequency would equal to $\Pr(x)\Pr(\text{hawkish})$. The Pointwise Mutual Information (PMI, Church and Hanks [1990]) between the string of text x and the word "hawkish" is defined as:

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

Pointwise mutual information is a central concept in information theory. Given two elements, PMI is a log-ratio indicating the amount of information that it is possible to gather about one element of the message when the other is observed.[18] A measure of the relative degree of association between the string x and the word "dovish" can be computed accordingly, hence obtaining the degree of dovishness that we can infer for x. In order to obtain a measure of orientation we can consider the relative PMI between the two polar concepts and obtain a theoretical score of semantic orientation (SO) of string x, as obtained from the PMI as:[19]

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

The Internet represents a very large corpus of text from which it is possible to obtain empirical frequencies of each string of text in a statement, and the words "hawkish" and "dovish". We implement the information retrieval (IR) process through hits counts on the search engine Google.[20] The empirical (feasible) semantic orientation score obtained by information retrieval on the Google search engine is:

[17] Or other relevant dichotomies, which in our context might be: restrictive/accommodative; active/passive; etc.

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

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

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

$$SO(x) = \log\left(\frac{\text{hits}(x\,\&\,\text{hawkish}) * \text{hits}(\text{dovish})}{\text{hits}(x\,\&\,\text{hawkish}) * \text{hits}(\text{hawkish})}\right). \tag{3}$$

where hits(x) assigns the number of hits in the search of query x. The SO score is defined over $(-\infty, \infty)$ and is increasing in the degree of hawkishness of the string of text x. Computing the scores associated with the strings X and X', which we presented earlier in this section, is straightforward by implementing six searches in Google. For example, for the hawkish sentence X we obtain the positive score:

$$SO(X) = \log(268 * 198,000/(24 * 840,000)) = .98,$$

and the negative score:

$$SO(X') = \log(970 * 198,000/(389 * 840,000)) = -.53,$$

for the dovish sentence X'. This example is representative of how an unsupervised, automated algorithm such as the SO-PMI can approximate a subjective interpretation of a string of text along the hawkish–dovish dimension.[21] We implement the scheme (3) on each sentence of a FOMC statement as follows: We first apply to the text a Brill [1994] part-of-speech tagger, a natural language processing algorithm used to automatically classify words in the lexical categories of nouns, adjectives, verbs, adverbs, pronouns and coordinating conjunctions. Based on this classification, we then apply other automated routines to obtain groups of words—chunks—corresponding to either verbal, noun or adjectival phrases, and run searches on sub-sentences composed of three or more of these chunks.[22] After obtaining the scores for each sentence, we average them over all sentences to obtain a score for the statement.

To compare the results of the PMI score and the heuristic score we consider a discretization of the semantic orientation score based on values of the PMI score outside of the interquartile range, or:

[21]The Web evolves continuously over time. Pages are substituted and dropped from Google caches over time and its index algorithm is run everyday. This implies that searches executed over time, even on the same data center, may differ. We run all our searches both in April 2007 and in August 2007 and found a correlation across hits well above 80 percent, indicating a substantial degree of overall persistence (the meaning of words is persistent and does not change as fast as the Web over time).

[22]In the implementation of the score, to avoid division by 0, we follow Turney [2002] and Turney and Littman [2002] and add the quantity 0.01 to the hits count as a form of Laplace smoothing. We perform the analysis only on those strings x of text for which more than two hits for the sum of hits(x & hawkish) and hits(x & dovish) are found. We also only included searches longer than three words to exclude incidentals (for instance running a query for the sub-sentence "in any event," is not meaningful by itself in our context). We experimented with three, four, five, and six chunks and obtained the best fit employing the four and five chunks. We report the analysis performed on five chunks, although the results in the following sections are robust to alternative sub-sentence length. Employing searches directly on whole sentences, however, did report zero hits very frequently and resulted in large number missing observations and noisy measurement. See the discussion below of coverage ratios. The robustness results are available from the authors upon request. The automated searches and natural language processing of the texts is implemented in Python 2.5, using routines from Liu [2004] and Bird and Loper [2006]

$$DSO(x) = \begin{cases} 1 & \text{if } \mathrm{cdf}(SO(x)) > 1 - \tau; \\ -1 & \text{if } \mathrm{cdf}(SO(x)) < \tau; \\ 0 & \text{otherwise}; \end{cases} \tag{4}$$

where $\tau = 1/4$. Figure 2 shows the discrete semantic orientation measure together with the intended federal funds rate. The time series of the DSO score is similar to that of the heuristic score shown in Figure 1 with the notable exception of the period starting in early 2001 and ending in early 2002. The reason for this discrepancy becomes clear in Figure 3, which shows the heuristic score along with the DSO scores. The hollow circles in the Figure indicate that the scores were based on matches on the search engine queries *covering less* than half of the sentences in the corresponding FOMC statement, while the full circles were obtained from statements with higher coverage. In the period around the end of 2001 the absence of hits on the search engine induces substantial noise in the unsupervised score. The low coverage of the statement in that period clearly highlights the sensitivity of the discrete SO score to a lack of hits. In the next sections we, therefore, retest the empirical results for coverage-related issues by dropping low-coverage observations from the sample. The sub-sampled scores are referred to as "covered" in the remaining of the paper. In addition we also impute missing values for the "covered" scores, as discussed below in this Section.

The limited coverage of sections of the statements is an obvious drawback of the automated scoring approach. However, as shown in Figure 4, the search hits covered most sentences following the gap of 2001 and all sentences for statements released over the past few years. Based on these results and the increasingly more important role played by the Internet in developed societies, we expect coverage gaps to be quite unlikely in applications of the automated scores to future statements.

It is also interesting to note that the overall alignment in the DSO and heuristic score levels is quite high. The continuous, unbounded measure SO does not match the heuristic index as well (as we should expect absent the discretization) when both are expressed in levels, however, the SO and HI scores are quite close in differences, a particularly relevant filter in our high-frequency analysis, which makes use of regressions that include first differences, rather than levels, of the scores.

Another simple score of semantic orientation that we will employ extensively is an adaptation of the intuitive measure proposed by Dave, Lawrence, and Pennock [2003]. This approach focuses on the relative imbalance (RI) between conditional frequencies as a measure of relative association (i.e. if a text string x reveals relatively more information about being hawkish than being dovish). From the joint frequencies $\Pr(x \,\&\, \text{hawkish})$ and $\Pr(x \,\&\, \text{dovish})$ (approximated by the number of hits estimated from the Web) we define the Relative Imbalance score as:

$$RI(x) = \frac{\Pr(\text{hawkish}|x) - \Pr(\text{dovish}|x)}{\Pr(\text{hawkish}|x) + \Pr(\text{dovish}|x)}. \tag{5}$$

The score in (5) is bounded between -1 and 1, symmetric, continuous, and again increasing in the degree of hawkishness of x. Figure 6 shows how the RI score closely follows the HI score, when it is covered by enough hits. We will use this score extensively in the low frequency results. In addition, we will impute missing values for the "covered" RI score using one-step ahead forecasts from an AR(1) model, estimated by the Kalman filter, that includes the HI index as an exogenous variable.[23] Figure 6 shows the RI "covered" score with imputed values.

Finally, we also attempted to address the fact that certain sections of the statement may be the focus of more attention, or reveal more information, than others about future policy rate moves, and thus lead to larger price responses. As an additional check we employ a weighting procedure based on the number of hits for each sentence in the two weeks following the release of a given statement using the Google News Archive search engine. This weighting procedure is applied to all scoring schemes presented above, and the scores thus obtained are referred to as "weighted" in the rest of the paper.[24]

A detailed description of all algorithms used to create the scores is available in Appendix.

4 Data Construction

This section describes the data used in the empirical analysis. Our data start in May 1999, the date in which the Federal Open Market Committee begun releasing statements systematically after all policy meetings, and end in August 2007. The data set, thus, includes 70 FOMC statements, excluding the one from the unscheduled meeting of September 17, 2001, because of a lack of financial data during the days immediately following the September 11 terrorist attacks. The empirical analysis is conducted at an intra-day (high) frequency, as well as at a monthly and intermeeting (low) frequency.

In the intra-day analysis, we study price responses of Treasury securities and Eurodollar futures to changes in the target for the federal funds rate and the content of FOMC statements—as measured by our semantic scores—during narrow windows of time around FOMC announcements. In the low-frequency analysis, instead, we look at the forecasting power of FOMC statements for future policy rate moves, using univariate models as well as vector autoregression(VAR) models that include measures of inflation, employment and risk-free yields at different maturities.

The empirical strategy employed in the high-frequency analysis follows that employed in recent related literature (e.g. Fleming and Piazzesi [2005], Gürkaynak, Sack, and Swanson [2005b], and Boukus and Rosenberg [2006]). By studying the response of asset yields during narrow time windows around the announcements, we isolate the impact of policy actions from other same-day events, such as economic data releases. We consider two temporal windows of different length: a "tight" window, which is thirty-minutes long—beginning ten minutes before and ending twenty minutes after the announcements—and a "wide" window one-hour long—starting -15 minutes and ending $+45$ minutes after. For each time window, our dependent variables include basis point changes of yields of on-the-run Treasury securities having maturities of 3 and 6-months, as well as 2-, 5-, 10-, and 30-years. We also consider short and medium-term Eurodollar futures contracts

[23] We considered alternative ARIMA models obtaining similar results.

[24] The Google News Archive is available at `news.google.com/archivesearch`.

which settle on forward rates on Eurodeposits.[25] We also construct empirical measures of the shape of the yield curve based on the selected quotes available for on-the-run Treasuries. Following Diebold, Rudebusch, and Boragan [2006] among others, we proxy: (i) the change in the level with the mean of the changes in the 3-, 2- and 10-year Treasuries; (ii) the change in the slope with the difference between the changes in the 10-year and the 3-month Treasury securities; (iii) the change in the curvature as the sum of the changes in the 2−year/3−month spread and in the 2−year/10−year spread.

Our explanatory variables include changes in policy rates and announcements. Following Kuttner [2001], we only include the component of the policy rate actions— hereafter, referred to as the monetary policy surprise — as measured by the scaled change in the current-month federal funds futures contract.[26] The change in the content of the policy announcement is, instead, measured as the change in the current semantic score relative to the score of the previous FOMC statement.

In the low-frequency analysis, data on the risk-free rates at different maturities are monthly averages from the daily estimates from off-the-run Treasuries of zero-coupon yields of Gurkaynak, Sack, and Wright [2007].

Table 1, 2 and 3 present summary statistics and correlation matrix for the data employed in the high-frequency analysis. All yields and monetary surprises are expressed in the tables as basis point changes. As shown in Table 3, the two automated scores display a high degree of correlation, and indeed, we find very similar results across the two measures in the high-frequency analysis. For an easier interpretation of the coefficients in Section 5 we standardize all semantic scores— imposing a zero mean and unit standard deviation—in order to make the regression coefficients easier to interpret and compare: The coefficients can all be interpreted as a basis point change of each dependent variable per unit standard deviation increase in each score.

5 High-Frequency Results

This section estimates the response of asset yields to changes in the content of FOMC statements, as measured by the semantic scores described above. In particular, we consider: (i) yields on Treasuries (Table 4); (ii) the level, slope and curvature of the Treasury yield curve (Table 8); (iii) yields of Eurodollar futures contracts (Table 10).

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

[26] Federal funds futures contract are priced on the the average effective federal funds rate for the month of expiration. The monetary policy surprise is calculated as: $(FF_{post} - FF_{pre}) * DM/(DM - d)$, where FF_{post} and FF_{pre} are the futures federal funds rate after and before the FOMC announcement, respectively. The scaling factor, $DM/(DM-d)$, adjusts for the averaging effect of the federal funds futures rates. See Kuttner [2001] for more details.

5.1 Specification

We follow a high-frequency identification approach that allows us to purge the response of asset yields to the policy action of the Fed from other news or events, such as economic data releases.[27] We present our empirical approach starting from Treasury securities. Let Δy_t^i be the change in the yield of security of maturity $i = 1, ..., m$ during the tight or wide time window around the FOMC announcement. Let MP_t and ΔS_t be, respectively, the monetary policy surprise at t and the change in the FOMC statement between periods t and $t-1$. [28] The two regressors differ in one important dimension. While certain components of policy actions and communication may be unexpected, others may be fully anticipated. This is fully accounted for by in the construction of the variable MP_t . Instead, ΔS_t includes both anticipated and unanticipated components, introducing a form of measurement error. Under reasonable expectation assumptions, this form of measurement error will bias the estimated coefficients toward zero. Hence, our quantitative estimates are better viewed as lower bounds of the effects of FOMC statements.

Although our methodology is close to the recent literature on the effects of FOMC announcements, our empirical setup is designed to allow cross-equation restrictions and tests on the coefficient vector β^i.[29] Define $\Delta X_t = [1 \;\; MP_t \;\; \Delta S_t]$. The specification expressed in stacked form can be written as:

$$\begin{bmatrix} \Delta y_t^1 \\ \Delta y_t^2 \\ \vdots \\ \Delta y_t^m \end{bmatrix} = \begin{bmatrix} \Delta X_t & 0 & \ldots & 0 \\ 0 & \Delta X_t & \ldots & 0 \\ & & \vdots & \\ 0 & 0 & \ldots & \Delta X_t \end{bmatrix} \begin{bmatrix} \beta^1 \\ \beta^2 \\ \vdots \\ \beta^m \end{bmatrix} + \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \\ \vdots \\ \varepsilon_t^m \end{bmatrix}, \tag{6}$$

which can be interpreted as a system of m seemingly unrelated equations.[30]

[27] The empirical validity of this approach is discussed in Cochrane and Piazzesi [2002] for one-day correlations and for intra-day Fleming and Piazzesi [2005].

[28] Our approach is flexible enough to extend to the analysis of statement transparency as well. We briefly explore this issue in Appendix B.

[29] For related work in the literature, see Fleming and Piazzesi [2005] and Gürkaynak, Sack, and Swanson [2005a], among others. See also Kuttner [2001] for the analysis of monetary policy surprises and Cook and Hahn [1989] for an application that does separate between expected and unexpected components of policies.

[30] We allow for a general within-announcement (t) covariance structure for the error terms across the m equations in the system:

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

within a clustered variance-covariance matrix:

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

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

5.2 Treasury Yields

Tables 4 and 5 report the regression results for changes in Treasury yields, respectively, over the "tight" and "wide" time windows around FOMC announcements. For a given linguistic score, each column shows the Treasuries' price responses to both monetary policy surprises and FOMC announcements. With the exclusion of the first column—a benchmark regression only containing the monetary policy surprise MP_t—the columns in the tables present model estimates containing a different linguistic score, ΔS_t.[31] The model specification, which is shown in (6), is a system of six equations containing as left-hand-side variable, yields on 3- and 6-month Treasuries, as well as 2-,5-,10-,30-year Treasuries. The rows of the Tables, show model estimates for the six simultaneous equations.

The first column of Table 4 reports the results for Treasuries' response to monetary policy surprise alone (MP_t) in the tight window. This specification is identical, up to sample coverage, to the the one considered in earlier literature, and the empirical findings are very similar.[32] We find evidence of a statistically significant effect of MP_t only on short-term yields with a substantial drop in the fraction of the variance explained by the surprise for medium- and longer-term yields.[33] A one-standard deviation increase in the monetary policy surprise in the tight window (8.37 basis points) produces positive and significant increases in the $3-$ and $6-$month bills and $2-$year note (respectively of 4.4, 3.8 and 2.8 basis points). As shown in the first column of Table 5, a one-standard deviation increase in the monetary policy surprise in the wide window (8.33 basis points) produces slightly larger increases in the the $3-$ and $6-$month bills and $2-$year note (respectively of 4.7, 4.2 and 3.0 basis points).

In columns 2 through 5 of Table 4 we include changes in stance of FOMC statements, ΔS_t, as measured by the different semantic scoring techniques. The coefficient on ΔS_t quantifies the effect of a change in the content of an FOMC statement, and changes in the scores are standardized to have a zero mean and a unit standard deviation, so that the coefficients express the basis point effect of a unit standard deviation increase in the corresponding score. Column 2 employs the change in the heuristic score as control. For assets with maturity above 6 months the column shows a significant and positive effect of the statement becoming more hawkish.[34] A one standard deviation increase produces an increase of 1.2 basis points for the $2-$year and 1.26 for the $5-$year note. Another intuitive way of interpreting the coefficients is as the effect of changing 1.7 sentences

[31] Given the novelty of our approach we feel compelled to provide evidence of its general robustness by showing several versions of the linguistic scores. The columns use in Table 5 and 4, in order: no score; heuristic score defined in (1); semantic orientation score defined in (3); discrete semantic orientation score defined in (4); relative imbalance score, defined in (5). The additional Tables 7 and 6 include the weighted and covered linguistic scores as discussed in Section 3. In particular the columns use, in order: weighted heuristic score; weighted semantic orientation score; weighted discrete semantic orientation score ; weighted relative imbalance score; covered semantic orientation score; covered discrete semantic orientation score; covered relative imbalance score.

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

[33] Notice however the non-monotonicity in the R^2 reported for the $30-$year yield, which presents a higher R^2 than the 2- and $5-$ year yields. This is mostly due to the inclusion in the sample of the unscheduled meetings of January 3, April 4, and September 17, 2001. All results below are robust to the exclusion of these three observations.

[34] For the tight sample. For the wide sample, Column 2 Table 5 shows an increase of one basis point for the $2-$year and 1.31 for the $5-$year yield.

from neutral to hawkish, given that the score HI belongs to $[-1, 1]$, it has a standard deviation of 0.32, and the average length of statement is 5.35 sentences.

In column 3 we introduce the semantic orientation score described in (3). The effect of a one standard deviation increase in the stance of the statement produces a humped effect along the yield curve with the positive effect first increasing and then decreasing with the yield maturity. The coefficient is statistically equal to zero for short-term yields, it peaks at 2 years (1.496 basis points with a robust clustered standard error of 0.59), and decreases again to about 1.08 for the 30−year bond. For all maturities above 2−years the estimates are very precise. Quantitatively the effects of the stance of the statements are large relative to the standard deviation of the dependent variables, as reported in Table 1, ranging from 3.77 (30−year maturity) basis points to 6.45 (2−year maturity). Substantially marked hump-shaped effects are present in the wide sample, as shown in Table 5. The coefficient on the 3−month yield is 0.55 (with a clustered standard error of 0.23), the coefficient on the 2−year note is 2.13 (with a clustered standard error of 0.73) and the coefficient on the 30−year bond is 0.90 (with a clustered standard error of 0.47).[35] Noticeably, the introduction of the semantic score increases substantially the explanatory power of the regression at medium- and long-term maturities relatively to the benchmark, which excludes the linguistic scores. The increase in R^2 is between 12 and 15 percent points in the wide sample.[36]

As shown in the bottom panel of Table 4 the p-value for the Wald tests of equality between the 3−month and 2−year coefficients reports a rejection at the 5 percent confidence level, indicating a stronger effect on medium-term yields, the opposite of what shown by the monetary policy surprise. The p-value for the Wald tests of equality between the 2−year and 30−year coefficients cannot reject the null of equality of the coefficients in the tight sample. In the wide sample, however, equality can be rejected at the 10 percent confidence level (Table 5). We therefore find evidence of excess sensitivity of very long-term yields to changes in FOMC communication, as such yields should theoretically display a much smaller, if any, response. Qualitatively the coefficients on the 10−year and 30−year bonds are substantially smaller than medium-term yields, but nonetheless often remain positive and significant. Interestingly widening the sampling window reduces this effect, suggesting that the excess price response tends to fade over time. We postpone additional discussion of this issue to Section 6.

Column 4 of Tables 4 and 5 employ the discrete version of the semantic orientation score as a measure of content of the statement. Again hump-shaped effects are present in the tight and wide

[35] We also performed our analysis with the 5-year and 10-year Treasury Inflation Protected Securities (TIPS) and the corresponding inflation compensation measures obtained by taking the difference between the on-the-run security and the corresponding TIPS. We found positive and significant effects of the statement stance on the implied real rates, but of negligible quantitative importance. TIPS are relatively illiquid assets and display variation around FOMC announcements two orders of magnitude lower than the corresponding on-the-run security (0.04 basis points). Due to the absence of meaningful quantitative variation we decided not to pursue this avenue further.

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

samples. For the former the coefficient on the 3−month bill is 0.42 (with a clustered standard error of 0.19), the coefficient on the 2−year note is 1.86 (with a clustered standard error of 0.57) and the coefficient on the 30− year bond is 1.28 (with a clustered standard error of 0.36). The wide sample presents quantitatively stronger effects in the middle range (2.22 basis points on 2−year and 5−year and rejecting equality between the 2−year and 30−year at 5 percent confidence). As for the case of HI the effects can be read in terms of number of sentences changing from neutral to hawkish. For instance the coefficient of 2.22 implies an effect of changing 1.65 sentences from neutral to hawkish, given $DSO \in [-1, 1]$ with a standard deviation of 0.31.

Column 5 of the tables employ a different unsupervised measure of stance, the relative imbalance score. This score, although based on different theoretical grounds, performs similarly in the analysis, both in terms of quantitative estimates and in terms of portion of the variance explained. Again on both dimensions the results indicate that the stance of FOMC statements has a significant economic and statistical effect with the expected signs. Interestingly across all columns of Table Tables 4 and 5, both for the tight and wide samples, the size of the coefficients on MP_t remain considerably stable, confirming the quantitative estimates of column (1). In order to check that the regression results are not just the artifact of few well-aligned outliers in the semantic orientation score or dependent variables, we report conditional scatter plot of yields for tight and wide samples in Figures 7 and 8. The partial regressions display the expected positive relationship between hawkishness and asset yields. We notice the particularly steep regression line for the 2-year note. Moreover we do not find prima facie evidence of our results being driven by outliers in the sample. Finally, as robustness checks complementing Tables 4 and 5 , we report in Tables 6 and 7, the same specifications considered so far employing both weighted and covered linguistic scores. The results of the weighted indices show substantially similar quantitative results to Table 2 with increases in precision of the estimates. The regressions employing covered scores are qualitatively similar to those reported in Tables 4 and 5, although they are not easily comparable in quantitative sense, because of the loss of about a one-third of the sample for lacking coverage. As mentioned above, the loss is concentrated in the early part of the sample and particularly in 2001 where the number of Google hits employed for estimating empirical frequencies is at times very low.

5.3 Shape of the Yield Curve

Tables 8 and 9 reports the results for the empirical proxies of level, slope, and curvature of the yield curve. These summary measures of the term structure depend deterministically on the yields employed as independent variables in the previous subsection. Hence the following results can be interpreted as linear combinations of the yield equations with appropriate constraints imposed on the coefficients. The advantage of focusing on level, slope, and curvature is that the relevant information concerning changes in the shape of the yield curve is condensed in simple t-tests. Anticipating the results below, we find that an increase in the hawkish stance of the FOMC announcement produces an increase in the level of the yield curve (as we have shown in Tables 4 and 5, the positive response is across all maturities). The stronger response on the long end relative to the short end (also documented in Tables 4 and 5) produces an increase in the slope of the yield curve, making it steeper. Finally, the curvature of the yield curve increases (increasing concavity), given the larger effect of the announcement on the 2−year note relative to both the 3−month bill and the 10−year note.

As for Tables 4 and 5, every column in Tables 8 and 9 employs a different semantic measure of the statement (ΔS_t), except for column (1) which is the benchmark. On the row dimension the Table reports the three simultaneous equations estimated by (6). Beginning from the level equation in the tight sample, the coefficient on MP_t ranges from 0.296 (column 1) to 0.32 (column 4), implying a positive effect of 2.6 bx, per standard deviation of the monetary policy surprise (8.37 basis points).[37] The effect of ΔS_t ranges from 0.80 (column 2) to 1.19 basis points (column 4).

Concerning the slope of the yield curve, the negative effect (thus flattening the curve) of the monetary policy surprise is the result of the null effect of changes in short-term rates on the long end, a result that should hold under the expectations hypothesis. The effect of the FOMC announcement is the opposite (thus steepening the curve). We estimate a change in the spread between 10−year note and the 3−month bill between 0.76 and 1.02 basis points per standard deviation increase of the announcement stance across the different measures. Quantitatively the size of the effects is substantial relative to the sample standard deviation of the spread for the tight sample (5.48 basis points). Finally notice that the automated indices generally provide more precise estimates relative to the heuristic score.

The change in the curvature (approximated as the sum of the changes in the 2−year/3−month spread and in the 2−year/10−year spread) ranges between 1.22 for HI and 2.01 basis points for DSO, all precisely estimated. Again the effects are economically significant if compared with a sample standard deviation of 7.22 basis points for the tight sample. In the wide sample the effects of the FOMC announcement on the slope and on the curvature occasionally become statistically weaker for specific communication measures, but overall confirm the results of the tight window sample. For all the level, slope and curvature equations there are sizeable increases in R^2 when controlling for ΔS_t.

5.4 Eurodollar Futures Contracts

We now report the regression results for four CME Eurodollar futures contracts in both the tight and wide samples. We study the first, fourth, sixth, and eight contracts, thus shedding light on the 3−month to 2−year maturity section of the yield curve.[38] Treasury yields are averages of future expected rates and term premia. By focusing on forward rates we are able to precisely time the effect of changes in the stance of the FOMC announcements on the term structure. We limit the analysis to the first eight contracts, which tend to be the most liquid around FOMC announcements. Table10 and 11 shows the different effects of FOMC stance across forward rates.

The first column, the benchmark, presents a positive effect of monetary surprise increases on

[37] For the remaining of this subsection we will refer to the tight sample unless otherwise mentioned.
[38] Results on all the first eight contracts are available from the authors upon request.

forward rates for every contract. Notice the decreasing coefficients on MP_t as contracts move further in the future, ranging from 0.641 to 0.396 basis points (in the tight sample and all statistically significant at 1 percent confidence). In columns 1 through 5 different linguistic indices are included. Again we generally find remarkably precise effects of the announcement, all going in the direction of higher expected interest rates when the stance of the statement becomes more hawkish. Particularly, excluding the first contract, the estimated impact of a one standard deviation increase in hawkishness is always positive and statistically significant for the unsupervised indices (both tight and wide sample). Quantitatively the effects are sizable. Consider for instance the discrete semantic score in column 3 where we find for the fourth contract a 2.68 basis points effect (per standard deviation), for the sixth contract a 2.56 basis points effect and 2.56 for the eight contract vis-a'-vis a standard deviation of the left-hand side between 7.51 and 8.49. The same specification in the wide window provides more (qualitative) information about the timing of the effect. The fourth contract shows a 2.73 basis points effect (per standard deviation), for the sixth contract a 4.03 basis points effect and 3.68 for the eight contract. Although confidence intervals are not sufficiently tight to pin down the peak of the FOMC announcement around 18 months, the qualitative evidence seems to point in this direction. We further address the issue of timing and the transmission mechanism in the low-frequency results below. At the bottom of Tables 10 and 11 , the Wald tests for equality between the ΔS_t coefficient of the first contract and the eight always strongly rejects the null in all specifications.

The addition of the stance indices produces a substantial increment in the portion of the variance explained by the specification in all columns and samples for all, but the first, Eurodollar contracts. For instance, by introducing the discrete semantic orientation score in the eight contract equation (column 3) in the tight sample, the R^2 of the regression increases to 24 percent from 11 percent in the benchmark including only the monetary policy surprise.

We now proceed to the analysis of the announcement effects at low frequency.

6 Low-Frequency Results

This section studies the link between central bank communication—as measured by the linguistics scores—, policy rate decisions, as well as measures of inflation, economic activity and nominal risk-free yields. We consider three empirical models: (i) a univariate model used to predict the level of the federal funds rate at different forecasting horizons using the linguistic scores, (ii) a univariate Taylor [1993]-type specification for the federal funds rate, and (ii) a vector autoregression (VAR) specification.[39]

[39] Although, the multivariate specification allows for feedback-effects between the variables, we find it easier and more direct to modify the univariate model specifications to include real-time data and future implied rates

6.1 Forecasting the Short-term Rates with the Automated Scores

This section evaluates the (in-sample) forecasting performance of the semantic orientation scores for future short-term rates. For the sake of parsimony, we restrict attention to the relative imbalance (RI) score.[40] We first analyze whether a more hawkish FOMC statement predicts higher short-term rates, as postulated in previous sections, by estimating the unconditional regression model:

$$r_{t+n} = \alpha^n + \beta^n RI_t + \varepsilon^n_{t+n} \tag{7}$$

where for every FOMC meeting t, r_{t+n} denotes the monthly average of the federal funds rate n-months after date t, as well as future 3-month Libor rates. We include in this analysis Libor rates in addition to the policy rate for reasons discussed below. The coefficient β^n measures the unconditional in-sample forecasting power of the linguistic score for rates n-periods after the meeting. Second, we analyze whether the content of the FOMC statement at date t includes relevant information to predict short-term rates beyond what already known to market participants fifteen-minutes ahead of the policy announcement—as implied by quotes on federal funds rate futures—, and in addition to the information implied by the surprise component of the announced policy action. In particular, we estimate the model:

$$r_{t+n} = \alpha^n + \beta^n RI_t + \gamma^n f^n_{t^-} + \lambda^n MP_t + \delta^n X_t + \varepsilon^n_{t+n} \tag{8}$$

where $f^n_{t^-}$ is the expected short-term rate at $t+n$ implied quotes on the corresponding future contract delivering at month $t+n$, fifteen minutes before the release of the policy announcement (t^-).[41] If the rates implied by futures quotes were efficient forecasts of futures rate, then they would incorporate all information available to the market up to the policy announcement. However, such efficiency might not always hold in the data due to, for example, (possibly time-varying) liquidity- and term-premia especially for futures contracts that settle after the first few months. In addition to the monetary policy surprise, MP_t, we therefore include in the regression a set of controls X_t that contain: (i) the most recent release of the PCE inflation rate (twelve-months change); (ii) the most recent reading of non-farm payroll employment (twelve months change). In addition we also include the lagged level of the score, RI_{t-1}, as a control for the expected component of the statement released at date t.

In (8) the coefficient β^n measures the response of short term rates n-months after the meeting

[40] As discussed in Section 3, the absence of hits on the search engine in the period around the end of 2001 induces substantial noise in our linguistics scores. In the analysis, we therefore always guarantee that the RI score has an appropriate coverage (at least 50 percent of the statement). Whenever missing values are present, we impute for the score using one-step ahead forecasts from an AR(1) model, estimated by the Kalman filter, that includes the HI index as an exogenous variable. Note that relative to the high-frequency analysis, it is substantially more important to capture the 2001 US recession in order to correctly identify the parameters for the low-frequency analysis, where the scores enter the empirical models in levels. We obtained similar low frequency results when using the DSO score.

[41] For example, for the month after the meeting, r_{t+1} is the realized average federal funds, whereas $f^1_{t^-}$ is the future rate implied by the quote on the federal funds rate contract that settles on the average realized effective federal funds rate for the same month.

to the content of the FOMC statement.[42]

Because the liquidity of the federal funds rate futures drops sharply for contracts expiring after the first few months–and thus their efficiency at estimating futures rates—we also consider Libor rates and implied rates from Eurodollar futures in models (7) and (8) to analyze the predictive power of the linguistic score for longer time horizons (up to two years out).[43] Although the Libor rate typically lies above the corresponding term federal funds rates, the spread between the two fluctuates within a fairly narrow band, so that Libor rates track federal funds rate fairly well.[44]

In Table 12, we report the estimates of models (7) and (8) for the federal funds rate, whereas Table 13 reports analogous estimates for the Libor rate. We estimate the model by maximum likelihood with heteroskedastic- and autocorrelation-consistent Newey and West [1987] standard errors with eight-lags truncation (corresponding to about one year). The estimates of the coefficients on β^n in (7) are positive and significant for all forecasting horizons, up to the sixth eurodollar contract–i.e. about one-half year out—as shown in Tables 12 and 13. Thus, unconditionally, a more hawkish FOMC statement positively predicts future interest rates hikes, or in other words, the stance of the FOMC statement at date t contains relevant information to predict future policy rate moves.

The estimation results of (7), however, cannot rule out that the information included in the statement is also contained in other observables known to market participants ahead of the announcement. Indeed, the policy stance of the FOMC—and thus the statement—likely depends on several other measures, including current and expected measures of inflation and economic activity. By employing implied future rates right ahead of the announcements, $f^n_{t^-}$, we can partial-out such information. In addition, by controlling for the monetary policy surprise, model (8) also separates the information included in the statement from what implied by the policy rate action.

As shown in the Tables, the magnitude of the coefficients β^n falls, as expected, when including these additional controls, but remains economically and statistically significant at conventional levels even four quarters out, for the Libor rate model. These results indicate that the statement contains information, beyond what already known to market participants, to predict short-term future rates up to a time horizon of one year. Quantitatively, the economic effects are reasonable, and are of the same order as those obtained using the VAR model specification, which is discussed in the next section. In panel a) the effect of a one-standard deviation increase in the stance of RI (0.23) on the federal funds rates ranges from 4.2 basis points month out to 15.7 five-months out. In panel b) the effects are 9.1 basis points one quarter out, 14.6 basis points at two quarters, 26.5 basis points at three quarters, 27.8 basis points at four quarters and 9.6 basis points at five

[42] The results that follow are best understood in the Granger sense.

[43] We only consider quarterly Eurodollar futures contracts. See footnote 25 for additional details on these futures. The left hand side rate included in (8) is, for each contract, the corresponding 3-month Libor rate on the settlement day of the contract.

[44] The spreads between the two rates tend to raise on relatively rare events corresponding to pressures in the interbank funding markets (e.g. ahead of year end in 1999 (Y2K) and 2007). For additional discussion about the use of Libor versus federal funds rates in computing market based monetary policy expectations, see Gürkaynak, Sack, and Swanson [2006].

quarters. The effects on the 3-month Libor rate are hump-shaped with a peak around three or four quarters out.

In conclusion, we find robust evidence of in-sample forecasting power of the automated scores for future policy rate moves, after partialling out for all information available to market participants ahead of the policy announcement and the monetary policy action. Similar results are also be shown in Section 6.3, which considers a multivariate (VAR) model specification that includes the linguistic scores and policy rates, as well as measures of inflation and economic activity.

6.2 Taylor Rule and Automated Scores

This section's goal is to assess the type of information contained in the FOMC statements as measured by our linguistic scores. In this analysis, we first estimate a Taylor rule for the federal funds rate, and then decompose the realized rate in two orthogonal components: the unsystematic component—the Taylor rule's residual or interest rate gap—and the systematic component, which we will refer to as the Taylor rule rate.[45] Finally, we compute cross-correlations between the two components and the linguistic scores at different leads and lags.[46]

This approach lets us establish whether our communication measures hold a stronger correlation with either of the two components. A higher correlation with the systematic component, would imply that the information content of the scores mostly relates to movements in rates due to changes in measures of inflation and economic activity; a higher correlation of the semantic scores with the unexplained component of the rule, instead, would indicate that the information in the linguistic scores refers to other factors affecting policy decisions, or a possible shift in the weights given by the FOMC to its goals of price stability and economic growth. We find below that communication appears to be related both to the current residual component of policy and to the future systematic component. The statement displays the highest correlation with Taylor rule's residuals almost contemporaneously , while it displays the highest correlation with the systematic component of policy with a lead of about seven meetings (about three quarters).

In the analysis we concentrate on real-time data at an intermeeting frequency (i.e. the data vintage available at the time of the FOMC meeting), in order to account for informational delays and corrections, which would otherwise affect both parameter estimates and residuals.[47]

We consider a Taylor-rule specification that incorporates partial interest rate adjustment—i.e. a lagged interest rate term—and serially correlated error terms, in order to account for the observed sluggishness of policy rates that has been highlighted in previous literature.[48] The model specification that we estimate between September 1987 and August 2007 is:

[45] As discussed below the Taylor rules include the lagged interest rate, but this term is not included in what we define as the Taylor rule rate

[46] The two-step procedure—rather than a joint estimation—allows us to decompose the policy rates in our sample using parameter estimates for the Taylor rule that include the full tenure of Chairman Greenspan.

[47] For a discussion of the role of real-time data in the study of policy rules see Orphanides [2001].

[48] See e.g. Clarida, Galí, and Gertler [2000]. For a discussion of partial adjustment and serially correlated error terms see Rudebusch [2002] and English, Nelson, and Sack [2003]. We experimented with alternative Taylor-rule specifications, and chose this specification based on a superior in-sample fit.

$$
\begin{aligned}
i_t &= \alpha i_{t-1} + (1-\alpha) i_t^T + \varepsilon_t && (9)\\
i_t^T &\equiv \beta_0 + \beta_\pi \pi_t + \beta_\pi \left(y_t - y_t^*\right), && (10)\\
\varepsilon_t &= \rho \varepsilon_{t-1} + v_t. && (11)
\end{aligned}
$$

where i_t^T is the Taylor rule rate, π_t is the inflation rate, and $y_t - y_t^*$ is the output gap. We measure inflation as the 12-month CPI inflation rate, and the output gap as the Greenbook-consistent measure of the output gap up to 2000 and the CBO output gap afterwards.[49] After estimating the Taylor rule, we obtain fitted values for the residual, $\hat{\varepsilon}$, and the Taylor rule rate, $\hat{i}_t^T$, and then compute cross-correlation with the Relative Imbalance score at different leads and lags.

The cross-correlation functions, $Corr(\hat{\varepsilon}_t, RI_{t-J},)$ $Corr(\hat{i}_t^T, RI_{t-J},)$, are shown in Figures 9 and 9 respectively. As shown in the graphs, the Relative Imbalance score displays relatively strong correlations both with the residual and the Taylor rule rate. In addition the linguistic score appears to have a relatively strong leading property with the Taylor rule rate of about seven-meetings (about three-quarters), and the score is almost contemporaneous to the Taylor residual—it leads by about one-meeting. In conclusion, our analysis suggests that the semantic score appears to contain information regarding both components of policy.

6.3 Vector Autoregression Analysis

The VAR models considered in this section include two monetary policy instruments: a policy interest rate and central bank's communication about future movements in the policy rate. The models are estimated on U.S. data beginning in May 1999 and ending in August 2007. The policy rate is the effective federal funds rate, whereas, Federal Reserve's communication about the future path of the federal funds rate is measured through our linguistic scores: By construction, these scores capture the degree of "hawkishness" or "dovishness" of the statements and more hawkish statements precede hikes in the federal funds rate in the sample considered, as already shown in the univariate models presented in the previous sections. After presenting the model and the identification strategy for the monetary policy shocks, the Section turns to a discussion of the impulse responses and the forecast error variance decompositions for the variables of interest, to unexpected innovations in the two policy instruments.

We estimate five VAR models that feature identical measures of inflation and economic activity, the federal funds rate and the RI score.[50] The VAR models differ in the maturity of the risk-free nominal yield included in each model. 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 three-months core-PCE inflation rate and the three-months (percentage) change in non-farm payroll employment; $\mathbf{S}_t$ denotes the policy block composed of, in order, the semantic orientation score and the federal

[49] We use headline CPI because it is the policy relevant measure for most of the period of analysis.

[50] See footnote 40 for additional discussion about the use of this score in the low frequency analysis.

funds rate. Finally, R_t^i, is the (continuously compounded) zero-coupon yield at maturity $i = \{3-month,\, 6-month, 1-year,\, 3-year,\, 10-year\}$ as calculated by Gurkaynak, Sack, and Wright [2007] from off-the-run Treasury securities. We will refer to the variables $\mathbf{Z}_t = [\mathbf{X}_t, \mathbf{S}_t]'$ as the core variables of the models.

The VAR models identify monetary policy shocks using a recursiveness assumption: First, the inflation rate and the change in non-farm payroll employment, $\mathbf{X}_t$, do not respond contemporaneously to innovations in the policy block $\mathbf{S}_t$ and the yield R_t^i. Furthermore, within the policy block, $\mathbf{S}_t$, the semantic orientation score is ordered first, so that the federal funds rate responds immediately to innovations in the score. We find this ordering to be preferable to the alternative of ordering the federal funds rate first in $\mathbf{S}_t$, as it allows to interpret the RI score as measuring the current stance of monetary policy. Nevertheless, the key results of this Section do not not depend on the ordering of the two policy instruments within $\mathbf{S}_t$.[51]

Most of the VAR models considered in earlier literature to identify monetary policy shocks do not include information regarding the term structure of risk-free rates. 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, $\mathbf{Z}_t$, either contemporaneously nor with a lag. Instead, the yield R_t^i can respond contemporaneously to innovations in the core variables.[52] The structural form of the VAR models can be written as:

$$\mathbf{a}\,[\mathbf{X}_t, \mathbf{S}_t, R_t^i]' = \mathbf{A}\,[\mathbf{X}_t, \mathbf{S}_t, R_t^i]' + \sigma\,[\varepsilon_{\mathbf{t}}^{\mathbf{X}}, \varepsilon_{\mathbf{t}}^{\mathbf{S}}, \varepsilon_t^{R^i}]', \tag{12}$$

for $i = \{\text{3-month, 6-month, 1-year, 2-year, 10-year}\}$, where:

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} & 0 \\ \mathbf{a}_{21} & \mathbf{a}_{22} & 0 \\ \mathbf{a}_{31} & \mathbf{a}_{32} & 1 \end{bmatrix}, \; \mathbf{A}(L) = \begin{bmatrix} \mathbf{A}_{11}(L) & \mathbf{A}_{11}(L) & 0 \\ \mathbf{A}_{21}(L) & \mathbf{A}_{22}(L) & 0 \\ \mathbf{A}_{31}(L) & \mathbf{A}_{32}(L) & A_{33}(L) \end{bmatrix}, \tag{13}$$

and the matrix σ is diagonal. The diagonal terms in the matrices $\mathbf{a}_{11}$ and $\mathbf{a}_{22}$ of (13) are equal to one, and the innovations ε_t^X, ε_t^S and $\varepsilon_t^{R^i}$ in (12) are serially- and mutually-uncorrelated identically-distributed structural shocks. It is important to note that, because of the zero elements in (13), the structural shocks of the policy block ε_t^S do not depend on the maturity of the yield included in each VAR model, or, in other words, each model identifies the same monetary policy shocks.[53] Based on the Akaike Information Criterion(AIC), we include four lags of the relevant variables in the model specifications in (12),[54] and because of the zero-restrictions in (13), we estimate the

[51] We interpret innovations of both policy instruments as capturing non-systematic components of policies due to changes in policy preferences of the committee—e.g., because of a change in the composition of the voting members of the FOMC—, the imperfect observability of the state of the economy when policies are set and/or other factors.

[52] For a review of this literature, see Christiano, Eichenbaum, and Evans [1999] and Leeper, Sims, and Zha [1996]. The specification of the model that we consider closely resembles that of Evans and Marshall [1998]. For a model, which, instead, uses information on asset yields to identify monetary policy shocks, see, e.g., Piazzesi [2005].

[53] Note that, as for the case of the high-frequency analysis of Section 5 the VAR model does not impose absence of arbitrage opportunities across different maturities in calculating the yield responses, as in, for example, in Piazzesi [2005]. Although interesting, these restrictions are beyond the scope of this section.

[54] Ivanov and Kilian [2005] find that, for monetary models of the sort considered here, the AIC provides the most

parameters in (12) within a seemingly unrelated framework.

Figures 11 and 12 show responses of the core variables, $\mathbf{Z}_t$, to an unexpected one-standard deviation increase in the innovations, $\varepsilon_{\mathbf{t}}^{\mathbf{S}}$, corresponding to the semantic orientation score and the federal funds rate. Figures 13 and 14, instead, display the yields' responses to these shocks. All responses in the figures are absolute deviations in basis points from the unshocked values, with the exception of the linguistic score, which is measured as an absolute deviation in the (unscaled) units of the variable. In all figures, the shaded areas represent two-standard error bootstrapped confidence bands for the corresponding impulse response.[55] Before discussing the results, it is important to note that a positive innovation to both variables can be interpreted as contractionary monetary policy shocks but, whereas positive innovations to the federal funds rate directly feed into higher short term rates, a positive innovation to the linguistic score may affect short term rates only in that it indicates a more "hawkish" stance of monetary policy.

Consider the response to an unexpected positive shock to the semantic orientation score. As shown in the bottom right panel of Figure 11 the response of the federal funds rate is hump-shaped with a peak of about 20 basis points around seven months after that the shock occurs. The response of the semantic orientation score, instead, is monotonically decreasing and relatively short lived with the score returning to its pre-shock level after six months. The responses of both core inflation and employment growth to a shock in the semantic orientation score are in general not statistically different from zero. Following a shock to the linguistic score, the core inflation rate falls for the first ten months. The response of nonfarm payroll employment, which is quite persistent, is positive for the first few months, and it then becomes negative.

Figure 12 shows the response of the core variables to an unexpected (positive) innovation in the federal funds rate. As shown in the bottom panel, the response of the semantic orientation score switches sign few times and it is never statistically different from zero. The response of the federal funds rate to its own innovation is, instead, positive and quite persistent. The core-inflation rate and nonfarm payroll employment growth display counter intuitive responses, as both increase after a contractionary shock to the federal funds rate, although the confidence bands around both responses are fairly large. Because of the limited availability of the linguistic scores, our sample size, which only starts in 1999, is small relative to similar studies in the literature. It is therefore not too surprising that the responses of core inflation and employment growth to monetary policy shocks are imprecisely estimated and have shapes that do not closely resemble those obtained in previous literature, which has considered much longer data samples.[56] Even with a relatively short sample

accurate estimate of impulse responses in small sample and data observed at a monthly frequency.

[55]To obtain the confidence bands, we resample 1,000 times from the fitted residuals of (12). The confidence bands are then constructed as the point estimates of the impulse response coefficients plus/minus two standard deviations of the impulse response coefficients across the resampled datasets.

[56]The data sample analyzed in the literature that uses VAR models to identify monetary policy shocks, typically starts in the early sixties and ends in the late nineties. In the VAR analysis, we obtained similar results when using alternative measures of inflation— e.g. headline—or output, such as an index of industrial production. The inclusion of commodity price indices, which are often considered to solve the "price puzzle", also did not significantly affect the shape of the impulse responses. Finally, the responses of core inflation and employment growth to a shock to the federal funds rate were similar when excluding the semantic orientation score from the models.

size, however, the model estimates show that an unexpected positive innovation in the semantic orientation score, that is, an unexpectedly more hawkish FOMC statement, is followed by policy rate hikes in subsequent months.

We now turn to the forecast error variance decompositions of the core variables that describe the portion of the conditional k-step ahead forecast error variance of each variable that can be accounted for by the innovation in the semantic orientation score and the federal funds rate. The estimated variance decompositions are shown in Table 14 along with bootstrapped standard errors of the point estimates, which are reported in square brackets below each corresponding value. There are two columns in the Table for each variable that show the portion of variance accountable for by the semantic orientation score—RI in the Table—and the federal funds rate—FFR in the Table. As shown in the second-to-last column, the shock to the linguistic score accounts for a significant portion of the forecast error variance of the federal funds rate with a maximum of about 30 percent six months out, and for slightly smaller amounts at 3-months and 1-year. The variance of the federal funds rate accounted for by its own shock, instead, is monotonically decreasing, with a maximum of about 65 percent at three months, and about 40 percent one year-ahead. The variance of the semantic orientation score accounted by its own shock is also monotonically decreasing and is quite large, with a maximum of about 95 percent after three months and 80 percent one year-ahead. The innovation of the federal funds rate, instead, accounts for a negligible portion of the variance of the score at all forecasting horizons. The portion of the variance of the core inflation rate and of the nonfarm payroll employment growth rate that is accounted for by either of the two monetary policy shocks is approximately equal to 10 percent over the different forecast horizons, although the estimates are relatively imprecisely estimated.

Now consider the responses of the risk-free nominal yields to unexpected shocks to the semantic orientation score and the federal funds rate. As shown in Figure 13, yields at all maturities increase on impact after a positive shock to the semantic orientation score, with the magnitude of the increase being smaller for yields of longer maturities. The responses of the 3-month and 6-month yields are hump-shaped with a peak response of about 15 and 10 basis points, respectively, about six months after that the shock occurs; both responses are statistically different from zero for about 8 months after the shock. The responses of the yields with maturities above 2-years are, instead, never statistically different from zero. Now consider the yields' responses to a federal funds rate shock, which are shown in Figure 13. As for the semantic orientation score, the impact responses of the yields are larger for yields with shorter maturities; in addition the responses of yields for maturities above five years are negative, albeit not significant. Relative to shocks to the linguistic score, the impact responses of yields to federal funds rate shocks are larger for shorter maturities and smaller at longer maturities. The 3-month and 6-month yields' responses are hump-shaped, with a peak of about 20 basis points at six months and 15 basis points around twelve months. The responses of the yields above 2-years are never statistically different from zero, and for yields of maturity of 5-years and above are negative for few months.

In sum short-term yields increase— on impact and for about one-and-half year afterward—

following both a semantic orientation score shock and a federal funds rate shock. The sign of the responses for yields of longer maturities are, instead, mixed but, compared with federal funds rate shocks, the yields' responses to the shocks to the linguistic scores are larger.

The portion of the forecast error variance of the yields accounted for by the linguistic score and the federal funds rate shocks are shown in Table 15. The two shocks account for a relatively small portion of the forecast error variance of yields with maturities beyond 2-years, while both shocks explain a significant portion of the forecast variance for the 3-months and 6-months yields at all horizons.

7 Conclusions

In this paper we develop a novel approach to measure the stance, content, and intensity of central bank's communication regarding future policy rate moves. In particular, we define an operative measure of the stance of FOMC statements along a "hawkishness" metric, based on the Committee's declared assessment of the risk to its long run goals of "price stability and maximum sustainable employment."

We follow a dual approach in the empirical analysis. We first rely on high-frequency (intra-day) identification around policy announcements in order to pin down the immediate response of the term structure of nominal risk-free interest rates to changes in FOMC statements, and we find this response to be significant in economic terms. A one standard deviation increase in the degree of hawkishness of a statement produces an increase of 2 basis points in 2-year and 5-year yields in a half-hour window around the FOMC announcement. Effects on the short-term yields (3- and 6-month bills) are quantitatively null and the point estimates are statistically different the medium-term yields. Effects on 10−year and 30− year yields are significant and generally half the size of the 2-year and 5-year Treasuries, although at times statistically indistinguishable.

Using data at low-frequency, we find in a VAR model that a one standard deviation unexpected "shock" to the policy stance of the statements yields to higher future policy rates, with a peak of about 20 basis point after about six months. Parameter estimates of a univariate model in addition show that the content of the FOMC statements has predictive power for the level of the federal funds rates up to eight quarters out. Finally, we find that FOMC statements contain information regarding both the systematic and unsystematic component of policy rate decisions, as decomposed by Taylor rules.

The new measures proposed in the paper rely on a branch of research in computer science and natural language processing that has been almost untapped by economists. The approach of this paper, in particular, has the advantage of being unsupervised, intuitive, and replicable across researchers. Investigation of these association measures and of the linguistic scores can be, for example, applied in the empirical validation of recent theoretical contributions aiming at understanding the role of communication in the interaction between economic agents. Applications outside policy announcements may range from the orientation of political campaigns to corporate earnings announcements.

A Appendix

A.1 Algorithm 1 Sentence Stance Score (Relative Imbalance): RI(Sentence)

1: RI = 0
  2: Split Sentence in n-grams
  3: N count all n-grams x contained in Sentence
  4: for all n-grams x contained in Sentence do
    5: let f_H and f_D be frequencies of (x AND hawkish) and (x AND dovish)
    6: run Internet searches (x AND hawkish) and (x AND dovish)
    7: generate the relative imbalance (f_H - f_D)/(f_H + f_D)
    8: RI = RI + (f_H - f_D)/(f_H + f_D)*1/N
  9: Return - RI

A.2 Algorithm 2 Statement Stance Score (Relative Imbalance): RI(Statement)

1: RIST = 0
  2: S count all sentences contained in Statement
  3: for each Sentence do
    4: s = RI(Sentence)
    5: RIST = RIST + s*1/S
  7: Return - RIST
  Note: Same procedure applies to all semantic orientation scores.

A.3 Algorithm 3 Heuristic Score: HI(Statement)

1: HIST = 0
  2: S count all sentences contained in Statement
  3: for each Sentence do
    4: s = heuristic mark of Sentence
    5: HIST = HIST + s*1/S
  7: Return - HIST

Note: the heuristic marks were assigned by the authors and one research assistants and then compared. The correlation between scorers was on average above 70 percent.

B Appendix: Transparency

We briefly discuss here the issue of transparency of the statement in terms of the precision of its message. Transparency of communication, like the average stance discussed above, is a difficult concept to implement empirically. However, one could consider the standard deviation of the stance of the sentences (or their subsections) within a statement as a proxy for how contradictory, difficult to interpret, complex or opaque a statement is. Consider a statement consisting of two sentences both being neutral versus a statement containing two apparently contradictory sentences (say, one dovish and one hawkish). To a first approximation, it seems reasonable to assume that the market participants may have a clearer view of the priorities of the Commitee and of future policy rate move in the former case relative to the latter.

Consistently with this approach, we measured the precision of the statement by taking the standard deviation of each sentence's semantic score. This measure of precision was then added to specification (6) directly as a control and interacted with ΔS_t. The evidence turned out inconclusive. In results which for parsimony of space we do not report,[57] we did not find systematic evidence of an increase in the effect of a change in the stance of the statement when the statement became more precise (i.e. lower within-statement standard deviation). Our conjecture is that standard deviations of sentence-level data, by compounding measurement error, further reduce the signal-to-noise ratio in the data (as opposed to averages, which tend to cancel out noise). We find (deliberate[58] or unvoluntary) transparency of the announcement an interesting avenue of future research, but we decided not to explore this further in the paper.

[57] All results are available from the authors upon request.

[58] Chairman Greenspan's Fedspeak was by his own account "a form of syntax destruction".

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

——— (2004b): "The Great Moderation," Speech to the meeting of the Eastern Economic Association, Washington, D.C., February 20, 2004.

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

BIRD, S., AND E. LOPER (2006): "Natural Language Toolkit," Available at: `nltk.sourceforge.net`.

BOUKUS, E., AND J. V. ROSENBERG (2006): "The information content of FOMC minutes," mimeo, Federal Reserve Bank of New York.

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

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

COOK, T., AND T. HAHN (1989): "The Effect of the Changes in the Federal Funds Target Rate on Market Interest Rates in the 1970s," *Journal of Monetary Economics*, 24, 331–51.

CRÉMER, J., L. GARICANO, AND A. PRAT (2006): "Language and the Theory of the Firm.," *Quarterly Journal of Economics*, 122(1), 373–407.

DAVE, K., S. LAWRENCE, AND D. M. PENNOCK (2003): "Mining the peanut gallery: opinion extraction and semantic classification of product reviews," *Proceedings of the 12th international conference on World Wide Web*, pp. 519–528.

DIEBOLD, F. X., G. D. RUDEBUSCH, AND A. S. BORAGAN (2006): "The macroeconomy and the yield curve: a dynamic latent factor approach," *Journal of Econometrics*, 131(1-2), 309–338.

ENGLISH, W. B., W. R. NELSON, AND B. P. SACK (2003): "Interpreting the Significance of the Lagged Interest Rate in Estimated Monetary Policy Rules," *Contributions to Macroeconomics*, 3(1), 1–16.

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.

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," *jme*, 54(4), 1132–1162.

GREIDER, W. (1989): *Secrets of the Temple: How the Federal Reserve Runs the Country.* Simon & Schuster.

GÜRKAYNAK, R. S., B. SACK, AND E. SWANSON (2005a): "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.

——— (2005b): "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models," *American Economic Review*, 95(1), 425–436.

——— (2006): "Market-based measures of monetary policy expectations," Working Paper Series 2006-04, Federal Reserve Bank of San Francisco.

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.

HAUSMAN, J., AND J. WONGSWAN (2006): "Global asset prices and FOMC announcements," International Finance Discussion Papers 886, Board of Governors of the Federal Reserve System (U.S.).

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

Kohn, D. L., and B. P. Sack (2003): "Central bank talk: does it matter and why?," Finance and Economics Discussion Series 2003-55, Board of Governors of the Federal Reserve System (U.S.).

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

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

Leeper, E. M., C. A. Sims, and T. Zha (1996): "What Does Monetary Policy Do?," *Brookings Papers on Economic Activity*, 54(1996-2), 1–78.

Liu, H. (2004): "MontyLingua: An end-to-end natural language processor with common sense.," Available at: `web.media.mit.edu/ hugo/montylingua.`

Morris, S., and H. S. Shin (2005): "Central Bank Transparency and the Signal Value of Prices," *Brookings Papers on Economic Activity*, 72(2005-2), 1–66.

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*, 91(4), 964–985.

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

Romer, C. D., and D. H. Romer (2000): "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, 90(3), 429–457.

Rudebusch, G. D. (2002): "Term structure evidence on interest rate smoothing and monetary policy inertia," *Journal of Monetary Economics*, 49(6), 1161–1187.

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

Svensson, L. E. O. (2006): "Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-transparency, Not Con," *American Economic Review*, 96(1), 448–452.

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.

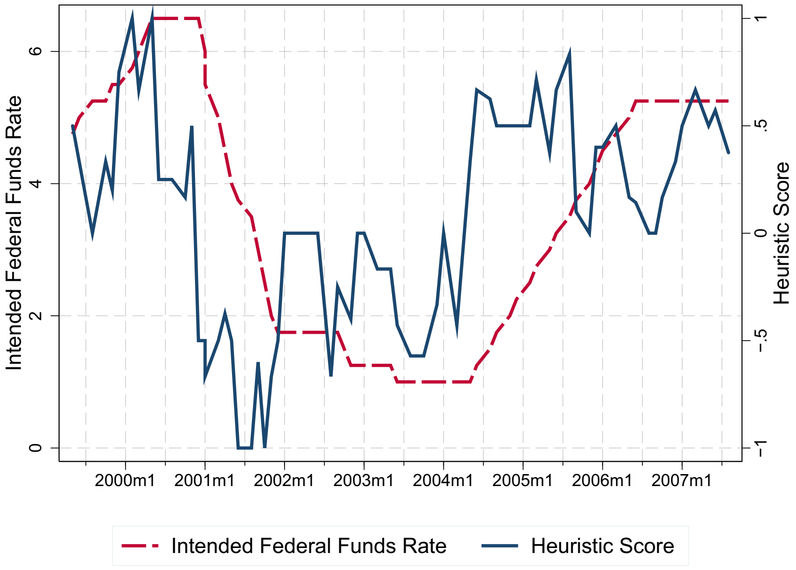
——— (2002): "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews," *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 417–424.

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.

——— (2003): "Measuring Praise and Criticism: Inference of Semantic Orientation from Association," *ACM Transactions on Information Systems*, 21, 315.

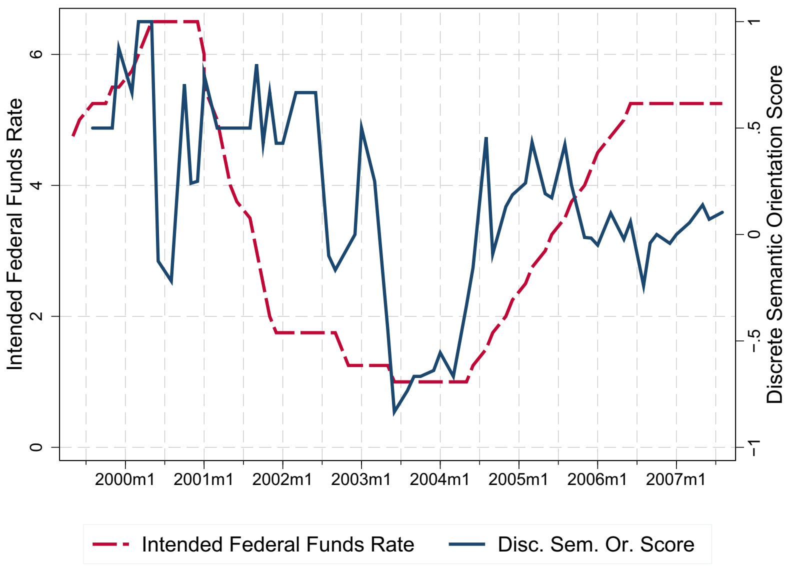
Woodford, M. (2005): "Central Bank Communication and Policy Effectiveness," NBER Working Papers 11898, National Bureau of Economic Research, Inc.

Figure 1: Heuristic Score and Intended Federal Funds Rate



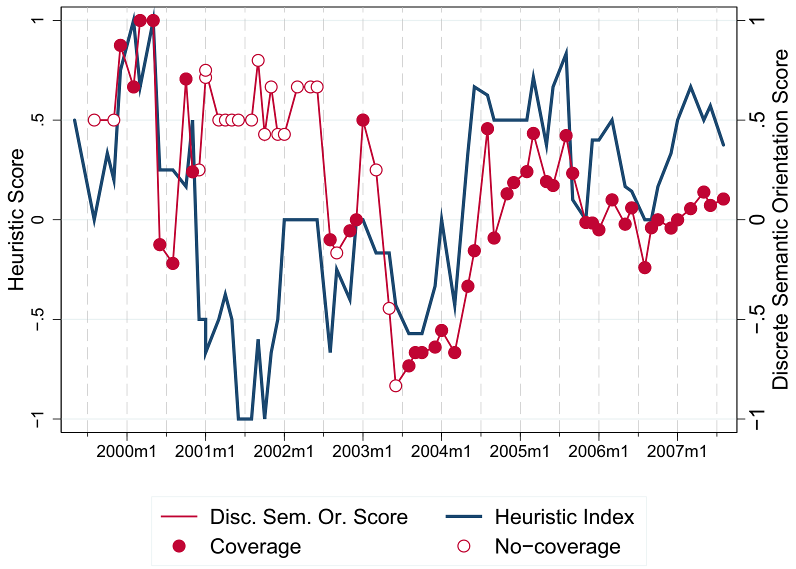
NOTES: Data at monthly frequency. The Heuristic Score is defined in Equation (1).

Figure 2: Discrete Semantic Orientation Score and Intended Federal Funds Rate



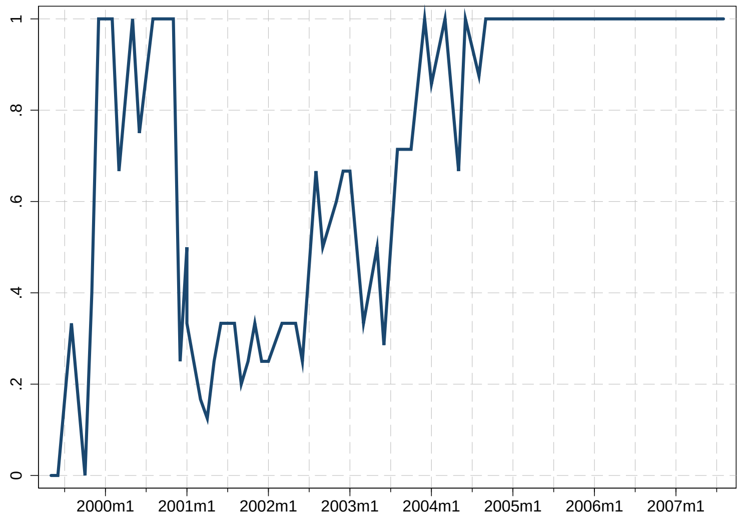
NOTES: Data at monthly frequency. The Discrete Semantic Orientation Score is defined in Equation (4).

Figure 3: Coverage of Discrete Semantic Orientation Score



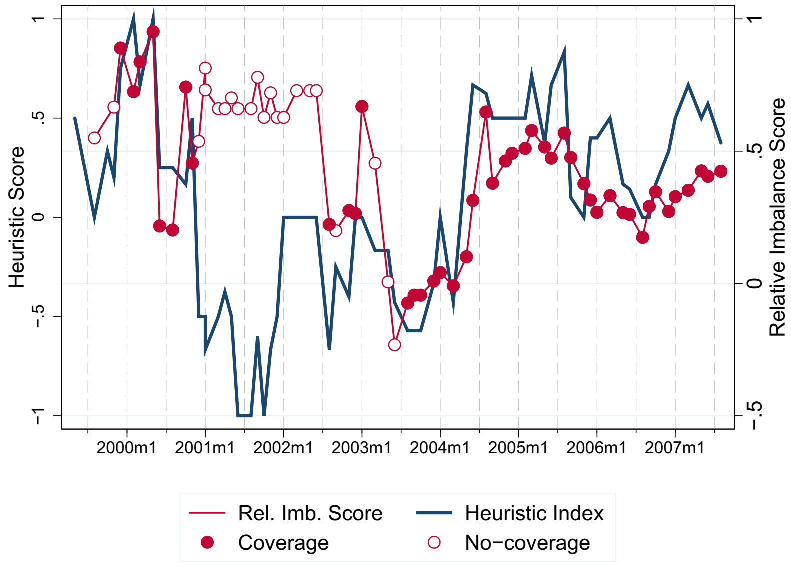
NOTES: Data at monthly frequency. The Discrete Semantic Orientation Score is defined in equation (4). The hollow circles indicate that the scores were based on matches on the search engine queries *covering less* than half the sentence in the FOMC statement, while the full circles were obtained from statements with a higher coverage.

Figure 4: Fraction of Sentences Covered by the Automated Scores



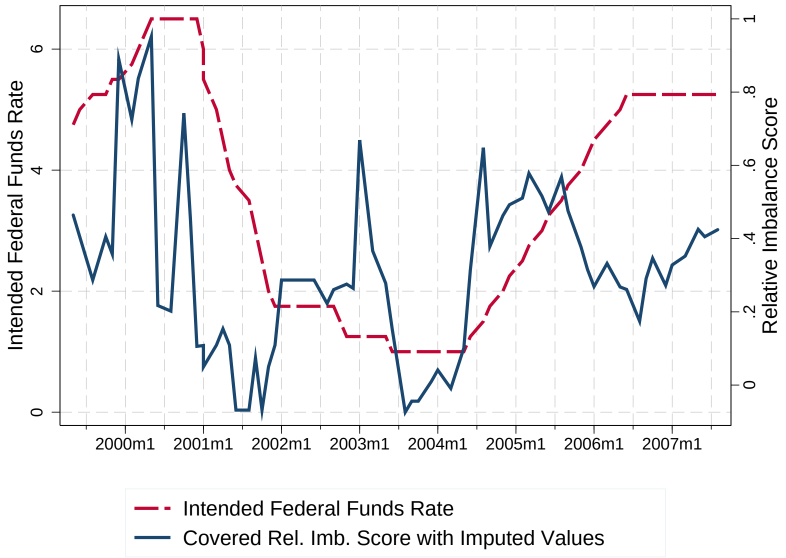
NOTES: Data at monthly frequency. The graph shows the fraction of sentences covered by search hits for each FOMC statement.

Figure 5: Coverage of Relative Imbalance Score



NOTES: Data at monthly frequency. Relative Imbalance Score is defined in equation (5). The hollow circles indicate that the scores were based on matches on the search engine queries *covering less* than half the sentence in the FOMC statement, while the full circles were obtained from statements with a higher coverage.

Figure 6: Covered Relative Imbalance Score with Imputed Values



NOTES: Data at monthly frequency. Relative Imbalance score is defined in equation (5). Whenever search hits cover less than half of the sentences in the statement, the score is imputed using one-step ahead forecasts from an AR(1) model, estimated by the Kalman filter, that includes the HI score as an exogenous variable.

Table 1: Summary Statistics for Dependent Variables in High-Frequency Analysis

Window	Tight						
Yield	3-month	6-month	2-year	5-year	10-year	30-year	
Stat.							
Mean	-1.20	-1.09	-0.47	0.27	0.15	0.28	
Median	0.00	0.00	0.00	0.00	-0.59	-0.20	
St. Dev.	4.77	4.86	6.45	5.71	4.39	3.78	
Min	-23.30	-24.30	-23.30	-11.30	-7.88	-6.50	
Max	9.00	8.00	21.55	22.88	16.16	14.00	
Window	Wide						
Yield	3-month	6-month	2-year	5-year	10-year	30-year	
Stat.							
Mean	-1.60	-1.33	-0.64	0.15	0.25	0.48	
Median	-1.00	0.00	0.40	0.00	0.20	0.11	
St. Dev.	5.27	5.00	7.52	6.73	5.20	4.09	
Min	-25.40	-25.40	-27.50	-19.00	-13.70	-8.30	
Max	11.00	9.00	18.24	18.42	13.61	14.00	
Window	Tight						
Variable	Level	Slope	Curvature	ED1	ED4	ED6	ED8
Stat.							
Mean	-0.51	1.35	0.12	-1.05	-1.38	-1.36	-1.04
Median	-0.47	0.31	1.03	-0.25	-1.50	-1.00	-1.00
St. Dev.	4.47	5.48	7.22	6.33	8.49	8.07	7.52
Min	-16.43	-6.88	-20.60	-38.00	-28.50	-24.00	-18.50
Max	13.24	20.80	24.94	10.50	24.50	25.00	25.50
Window	Wide						
Variable	Level	Slope	Curvature	ED1	ED4	ED6	ED8
Stat.							
Mean	-0.66	1.85	0.07	-0.99	-1.11	-1.71	-1.14
Median	-0.34	0.19	1.70	0.00	-0.50	0.50	0.00
St. Dev.	5.11	6.39	8.68	6.48	9.86	12.09	10.56
Min	-20.90	-10.01	-20.42	-35.00	-28.50	-63.50	-54.50
Max	11.28	30.00	21.33	12.00	24.50	26.00	24.00

NOTES: Basis points change during "tight" and "wide" time windows around FOMC announcements. Yields are for on-the-run Treasury securities. ED1-8 refer to the nearest-eighth quarterly Eurodollar futures contract. Level, slope and curvature are defined as: (i) mean of the 3-, 2- and 10-year Treasuries; (ii) difference between the 10-year and the 3-month Treasuries; (iii) sum of the changes in the 2−year/3−month spread and in the 2−year/10−year spread. Number of observations: 69.

Table 2: Summary Statistics for Independent Variables in High-Frequency Analysis

Variable	MP (Tight)	MP (Wide)	ΔRI	ΔSO	ΔHI
Mean	-1.33	-1.19	0.00	-0.03	0.00
Median	0.00	0.00	0.00	0.00	0.00
St. Dev.	8.37	8.33	0.19	1.09	0.32
Min	-43.75	-46.25	-0.73	-4.50	-1.00
Max	13.50	12.50	0.54	3.27	0.76
N. Obs.	69	69	65	65	69

NOTES: MP(Tight/Wide) is the monetary policy surprise during a "tight"/"wide" time window around FOMC announcements computed from the current-month federal funds futures contract. ΔRI is the change in the Relative Imbalance score defined in equation (5). ΔSO is the change in the Semantic Orientation score defined in equation (3). ΔHI is change in the Relative Imbalance score defined in equation (1).

Table 3: Pairwise Correlation Matrix for Independent Variables in High-Frequency Analysis

Variable	MP1t	MP1w	ΔRI	ΔSO	ΔHI
MP1t	1				
MP1w	1.00	1			
ΔRI	-0.19	-0.18	1		
ΔSO	-0.20	-0.20	0.91	1	
ΔHI	-0.12	-0.14	0.33	0.26	1

NOTES: For a definition of the variables see Table 1. Number of observations: 65.

Table 4: Regression Results for Treasury Yields: Tight Window

Δ Score:	Excl.	ΔHI	ΔSO	ΔDSO	ΔRI
Dependent Variable: Δ 3-month yield					
MP	0.568	0.576	0.584	0.583	0.583
	[0.052]***	[0.054]***	[0.051]***	[0.053]***	[0.053]***
Δ Score		0.462	0.545	0.542	0.544
		[0.256]*	[0.230]**	[0.231]**	[0.216]**
R^2	0.81	0.82	0.85	0.85	0.85
Dependent Variable: Δ 6-month yield					
MP	0.503	0.511	0.525	0.526	0.525
	[0.052]***	[0.053]***	[0.043]***	[0.047]***	[0.048]***
Δ Score		0.449	0.719	0.766	0.768
		[0.308]	[0.368]*	[0.361]**	[0.373]**
R^2	0.70	0.72	0.76	0.77	0.77
Dependent Variable: Δ 2-year yield					
MP	0.366	0.383	0.399	0.402	0.398
	[0.156]**	[0.161]**	[0.132]***	[0.138]***	[0.140]***
Δ Score		0.988	1.972	2.22	2.082
		[0.681]	[0.721]***	[0.721]***	[0.753]***
R^2	0.03	0.07	0.14	0.15	0.14
Dependent Variable: Δ 5-year yield					
MP	0.137	0.16	0.173	0.173	0.17
	[0.179]	[0.183]	[0.151]	[0.161]	[0.163]
Δ Score		1.305	2.126	2.224	2.132
		[0.612]**	[0.727]***	[0.664]***	[0.681]***
R^2	0.03	0.07	0.14	0.15	0.14
Dependent Variable: Δ 10-year yield					
MP	0.018	0.039	0.051	0.046	0.045
	[0.137]	[0.141]	[0.110]	[0.123]	[0.123]
Δ Score		1.151	1.921	1.776	1.812
		[0.496]**	[0.551]***	[0.476]***	[0.483]***
R^2	0.00	0.05	0.16	0.13	0.14
Dependent Variable: Δ 30-year yield					
MP	-0.111	-0.092	-0.101	-0.099	-0.101
	[0.091]	[0.095]	[0.077]	[0.083]	[0.084]
Δ Score		1.029	0.898	1.033	0.991
		[0.367]***	[0.472]*	[0.418]**	[0.396]**
R^2	0.05	0.11	0.12	0.14	0.13
Obs.	69	68	64	64	64
p-val. 3M=2Y		0.37	0.03	0.01	0.02
p-val. 2Y=30Y		0.94	0.06	0.02	0.04

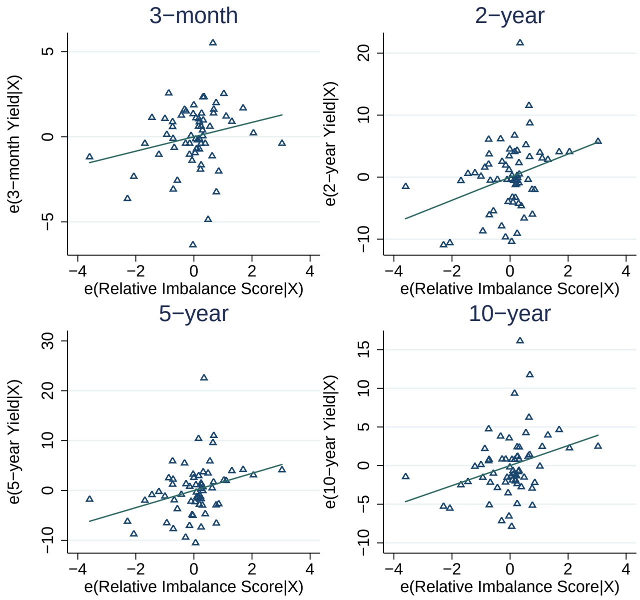
NOTES: *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 5: Regression Results for Treasury Yields: Wide Window

Δ Score:	Excl.	ΔHI	ΔSO	ΔDSO	ΔRI
Dependent Variable: Δ 3-month yield					
MP	0.523	0.526	0.53	0.531	0.53
	[0.054]***	[0.056]***	[0.050]***	[0.053]***	[0.053]***
Δ Score		0.169	0.349	0.423	0.389
		[0.214]	[0.210]*	[0.199]**	[0.186]**
R^2	0.84	0.85	0.86	0.86	0.86
Dependent Variable: Δ 6-month yield					
MP	0.453	0.459	0.468	0.476	0.474
	[0.070]***	[0.072]***	[0.063]***	[0.064]***	[0.064]***
Δ Score		0.392	0.439	0.77	0.729
		[0.394]	[0.394]	[0.486]	[0.500]
R^2	0.61	0.65	0.67	0.69	0.68
Dependent Variable: Δ 2-year yield					
MP	0.335	0.355	0.363	0.371	0.367
	[0.127]***	[0.132]***	[0.106]***	[0.112]***	[0.113]***
Δ Score		1.206	1.496	1.863	1.791
		[0.611]**	[0.594]**	[0.572]***	[0.581]***
R^2	0.03	0.08	0.10	0.13	0.12
Dependent Variable: Δ 5-year yield					
MP	0.12	0.14	0.149	0.155	0.151
	[0.103]	[0.107]	[0.082]*	[0.089]*	[0.090]*
Δ Score		1.256	1.416	1.718	1.616
		[0.588]**	[0.494]***	[0.456]***	[0.448]***
R^2	0.03	0.08	0.10	0.13	0.12
Dependent Variable: Δ 10-year yield					
MP	0.03	0.046	0.053	0.057	0.054
	[0.061]	[0.066]	[0.048]	[0.054]	[0.054]
Δ Score		1.018	1.111	1.296	1.273
		[0.494]**	[0.369]***	[0.351]***	[0.342]***
R^2	0.00	0.06	0.07	0.10	0.10
Dependent Variable: Δ 30-year yield					
MP	-0.123	-0.101	-0.103	-0.099	-0.103
	[0.090]	[0.095]	[0.073]	[0.080]	[0.081]
Δ Score		1.384	1.078	1.281	1.177
		[0.362]***	[0.453]**	[0.363]***	[0.324]***
R^2	0.07	0.21	0.17	0.20	0.19
Obs.	69	68	64	64	64
p-val. 3M=2Y		0.05	0.02	0.00	0.00
p-val. 2Y=30Y		0.72	0.53	0.39	0.38

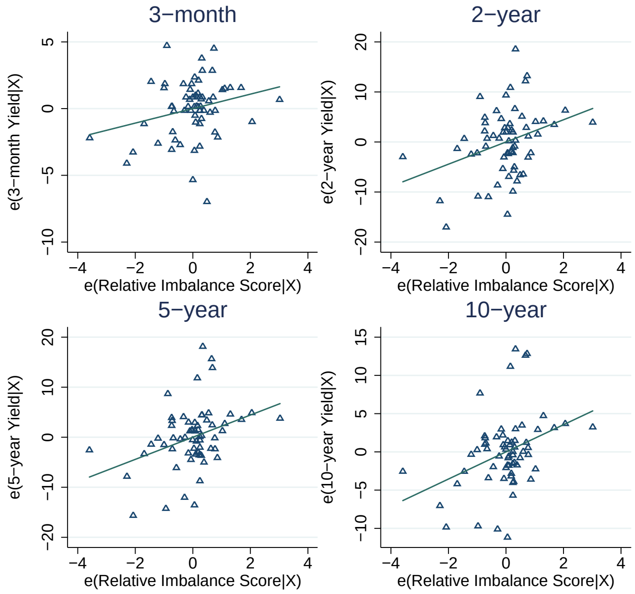
Notes: *** significant at 1%, ** significant at 5%, * significant at 10%.

Figure 7: Conditional Scatter Plots for Treasury Yields Regressions: Tight Window



Notes: The graphs show the conditional scatter plots for column (5) of Table 4

Figure 8: Conditional Scatter Plots for Treasury Yields Regressions: Wide Window



Notes: The graphs show the conditional scatter plots for column (5) of Table 5

Table 6: Regression Results for Treasury Yields: Robustness for Tight Window

Δ Score	ΔHI^W	ΔSO^W	ΔDSO^W	ΔRI^W	ΔSO^C	ΔDSO^C	ΔRI^C
Dependent Variable: Δ 3-month yield							
MP	0.523	0.531	0.529	0.528	0.47	0.486	0.477
	[0.055]***	[0.049]***	[0.053]***	[0.053]***	[0.099]***	[0.091]***	[0.093]***
Δ Score	0.293	0.287	0.277	0.243	0.159	0.173	0.183
	[0.178]*	[0.208]	[0.219]	[0.225]	[0.194]	[0.181]	[0.183]
R^2	0.85	0.85	0.85	0.85	0.41	0.41	0.41
Dependent Variable: Δ 6-month yield							
MP	0.452	0.469	0.476	0.474	0.213	0.216	0.212
	[0.072]***	[0.063]***	[0.066]***	[0.066]***	[0.098]**	[0.093]**	[0.094]**
Δ Score	0.636	0.336	0.654	0.583	0.017	-0.002	0.03
	[0.345]*	[0.390]	[0.489]	[0.509]	[0.161]	[0.215]	[0.177]
R^2	0.66	0.67	0.68	0.68	0.04	0.04	0.04
Dependent Variable: Δ 2-year yield							
MP	0.333	0.374	0.374	0.375	0.151	0.264	0.217
	[0.134]**	[0.102]***	[0.114]***	[0.114]***	[0.312]	[0.303]	[0.305]
Δ Score	1.684	1.475	1.716	1.708	1.041	1.054	1.066
	[0.600]***	[0.551]***	[0.580]***	[0.588]***	[0.240]***	[0.272]***	[0.254]***
R^2	0.13	0.09	0.10	0.10	0.03	0.04	0.05
Dependent Variable: Δ 5-year yield							
MP	0.118	0.159	0.155	0.155	0.035	0.134	0.082
	[0.109]	[0.080]**	[0.093]*	[0.094]*	[0.329]	[0.315]	[0.318]
Δ Score	1.69	1.367	1.464	1.429	1	1.105	1.143
	[0.616]***	[0.453]***	[0.476]***	[0.477]***	[0.205]***	[0.280]***	[0.262]***
R^2	0.13	0.09	0.10	0.10	0.03	0.04	0.05
Dependent Variable: Δ 10-year yield							
MP	0.028	0.062	0.058	0.059	-0.188	-0.089	-0.138
	[0.066]	[0.047]	[0.058]	[0.057]	[0.265]	[0.254]	[0.254]
Δ Score	1.308	1.128	1.144	1.152	0.973	1.05	1.087
	[0.490]***	[0.345]***	[0.386]***	[0.386]***	[0.245]***	[0.304]***	[0.292]***
R^2	0.10	0.07	0.08	0.08	0.04	0.05	0.06
Dependent Variable: Δ 30-year yield							
MP	-0.125	-0.095	-0.097	-0.099	-0.253	-0.142	-0.199
	[0.095]	[0.070]	[0.083]	[0.083]	[0.225]	[0.216]	[0.209]
Δ Score	1.303	1.068	1.143	1.06	1.053	1.1	1.171
	[0.354]***	[0.503]**	[0.392]***	[0.378]***	[0.512]**	[0.479]**	[0.448]***
R^2	0.20	0.17	0.18	0.17	0.07	0.09	0.10
Obs.	68	64	64	64	41	41	41
p-val. 3M=2Y	0.01	0.01	0.00	0.01	0.01	0.01	0.01
p-val. 2Y=30Y	0.39	0.50	0.37	0.33	0.99	0.94	0.86

NOTES: *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 7: Regression Results for Treasury Yields: Robustness for Wide Window

Δ Score:	ΔHI^W	ΔSO^W	ΔDSO^W	ΔRI^W	ΔSO^C	ΔDSO^C	ΔRI^C
Dependent Variable: Δ 3-month yield							
MP	0.568	0.585	0.581	0.581	0.463	0.507	0.49
	[0.053]***	[0.051]***	[0.053]***	[0.053]***	[0.194]**	[0.199]**	[0.196]**
Δ Score	0.674	0.455	0.396	0.374	0.421	0.378	0.43
	[0.213]***	[0.232]*	[0.251]	[0.251]	[0.240]*	[0.202]*	[0.192]**
R^2	0.83	0.84	0.84	0.84	0.35	0.34	0.35
Dependent Variable: Δ 6-month yield							
MP	0.503	0.53	0.525	0.526	0.479	0.493	0.487
	[0.054]***	[0.042]***	[0.049]***	[0.049]***	[0.064]***	[0.060]***	[0.063]***
Δ Score	0.692	0.706	0.654	0.66	0.18	0.211	0.213
	[0.264]***	[0.320]**	[0.332]**	[0.335]**	[0.126]	[0.134]	[0.149]
R^2	0.73	0.76	0.76	0.76	0.27	0.28	0.28
Dependent Variable: Δ 2-year yield							
MP	0.366	0.414	0.409	0.409	0.439	0.513	0.494
	[0.162]**	[0.125]***	[0.136]***	[0.137]***	[0.339]	[0.316]	[0.333]
Δ Score	1.724	1.995	2.192	2.098	0.946	1.122	1.01
	[0.649]***	[0.591]***	[0.657]***	[0.653]***	[0.277]***	[0.326]***	[0.358]***
R^2	0.11	0.14	0.13	0.13	0.04	0.06	0.06
Dependent Variable: Δ 5-year yield							
MP	0.137	0.189	0.176	0.177	0.057	0.154	0.116
	[0.186]	[0.145]	[0.164]	[0.164]	[0.254]	[0.242]	[0.247]
Δ Score	1.878	2.118	2.054	1.988	1.171	1.326	1.32
	[0.667]***	[0.618]***	[0.575]***	[0.565]***	[0.235]***	[0.347]***	[0.328]***
R^2	0.11	0.14	0.13	0.13	0.04	0.06	0.06
Dependent Variable: Δ 10-year yield							
MP	0.018	0.068	0.05	0.053	-0.051	0.056	0.015
	[0.142]	[0.102]	[0.124]	[0.123]	[0.189]	[0.200]	[0.186]
Δ Score	1.443	2.008	1.735	1.771	1.195	1.263	1.307
	[0.541]***	[0.490]***	[0.441]***	[0.437]***	[0.263]***	[0.348]***	[0.329]***
R^2	0.08	0.17	0.13	0.13	0.06	0.07	0.08
Dependent Variable: Δ 30-year yield							
MP	-0.111	-0.093	-0.095	-0.096	-0.189	-0.144	-0.169
	[0.095]	[0.073]	[0.084]	[0.084]	[0.181]	[0.182]	[0.174]
Δ Score	1.111	0.938	1.068	0.984	0.553	0.641	0.695
	[0.410]***	[0.496]*	[0.398]***	[0.378]***	[0.334]*	[0.385]*	[0.368]*
R^2	0.13	0.12	0.14	0.13	0.02	0.03	0.03
Obs.	68	64	64	64	41	41	41
p-val. 3M=2Y	0.06	0.01	0.00	0.01	0.10	0.04	0.13
p-val. 2Y=30Y	0.17	0.05	0.03	0.03	0.32	0.21	0.49

NOTES: *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 8: Regression Results for the Shape of the Yield Curve: Tight Window

Δ Score:		ΔHI	ΔSO	ΔDSO	ΔRI
Dependent Variable: Δ Yield Level					
MP	0.296	0.309	0.315	0.32	0.317
	[0.075]***	[0.079]***	[0.059]***	[0.066]***	[0.066]***
Δ Score		0.798	0.985	1.194	1.151
		[0.392]**	[0.360]***	[0.343]***	[0.341]***
R^2	0.31	0.36	0.39	0.41	0.41
Dependent Variable: Δ Yield Slope					
MP	-0.493	-0.479	-0.477	-0.475	-0.476
	[0.024]***	[0.024]***	[0.022]***	[0.021]***	[0.021]***
Δ Score		0.849	0.762	0.873	0.884
		[0.476]*	[0.282]***	[0.280]***	[0.276]***
R^2	0.57	0.61	0.62	0.63	0.63
Dependent Variable: Δ Yield Curvature					
MP	0.118	0.138	0.143	0.154	0.149
	[0.186]	[0.191]	[0.175]	[0.174]	[0.175]
Δ Score		1.224	1.531	2.007	1.921
		[0.733]*	[0.790]*	[0.767]***	[0.789]**
R^2	0.02	0.05	0.06	0.09	0.09
Obs.	69	68	64	64	64

NOTES: *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 9: Regression Results for the Shape of the Yield Curve: Wide Window

Δ Score:		ΔHI	ΔSO	ΔDSO	ΔRI
Dependent Variable: Δ Yield Level					
MP	0.317	0.333	0.344	0.344	0.342
	[0.103]***	[0.107]***	[0.082]***	[0.090]***	[0.091]***
Δ Score		0.867	1.48	1.513	1.48
		[0.426]**	[0.453]***	[0.435]***	[0.444]***
R^2	0.27	0.31	0.39	0.39	0.39
Dependent Variable: Δ Yield Slope					
MP	-0.55	-0.537	-0.533	-0.538	-0.538
	[0.130]***	[0.132]***	[0.119]***	[0.125]***	[0.125]***
Δ Score		0.689	1.376	1.234	1.268
		[0.462]	[0.500]***	[0.420]***	[0.430]***
R^2	0.51	0.53	0.60	0.59	0.59
Dependent Variable: Δ Yield Curvature					
MP	0.146	0.152	0.163	0.176	0.168
	[0.181]	[0.186]	[0.177]	[0.171]	[0.174]
Δ Score		0.363	1.478	2.122	1.807
		[0.866]	[0.899]	[0.934]**	[0.988]*
R^2	0.02	0.02	0.04	0.07	0.06
Obs.	69	68	64	64	64

NOTES: *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 10: Regression Results for Eurodollar Futures: Tight Window

Δ Score:	Excl.	ΔHI	ΔSO	ΔDSO	ΔRI
Dependent Variable: Δ ED1					
MP	0.641	0.646	0.659	0.664	0.66
	[0.131]***	[0.133]***	[0.120]***	[0.124]***	[0.125]***
Δ Score		0.314	0.832	1.046	0.943
		[0.304]	[0.456]*	[0.370]***	[0.368]**
R^2	0.7	0.7	0.8	0.8	0.8
Dependent Variable: Δ ED4					
MP	0.459	0.488	0.509	0.521	0.516
	[0.119]***	[0.125]***	[0.092]***	[0.098]***	[0.099]***
Δ Score		1.829	2.134	2.678	2.612
		[0.859]**	[0.821]***	[0.912]***	[0.950]***
R^2	0.2	0.3	0.3	0.3	0.3
Dependent Variable: Δ ED6					
MP	0.382	0.411	0.431	0.442	0.439
	[0.098]***	[0.103]***	[0.072]***	[0.077]***	[0.078]***
Δ Score		1.822	2.045	2.556	2.546
		[0.813]**	[0.767]***	[0.811]***	[0.843]***
R^2	0.2	0.2	0.2	0.3	0.3
Dependent Variable: Δ ED8					
MP	0.296	0.328	0.347	0.359	0.353
	[0.089]***	[0.097]***	[0.064]***	[0.071]***	[0.072]***
Δ Score		1.965	1.97	2.559	2.399
		[0.736]***	[0.743]***	[0.777]***	[0.789]***
R^2	0.109	0.185	0.192	0.239	0.226
Obs.	69	68	64	64	64
p-val. Q1=Q8		0.01	0.02	0.01	0.01

Notes: *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 11: Regression Results for Eurodollar Futures: Wide Window

Δ Score:	Excl.	ΔHI	ΔSO	ΔDSO	ΔRI
Dependent Variable: Δ ED1					
MP	0.625	0.632	0.645	0.648	0.644
	[0.106]***	[0.109]***	[0.093]***	[0.097]***	[0.099]***
Δ Score		0.377	0.966	1.136	0.974
		[0.416]	[0.445]**	[0.377]***	[0.373]***
R^2	0.6	0.7	0.7	0.7	0.7
Dependent Variable: Δ ED4					
MP	0.403	0.43	0.455	0.462	0.457
	[0.167]**	[0.172]**	[0.140]***	[0.145]***	[0.147]***
Δ Score		1.51	2.325	2.726	2.582
		[0.971]	[0.982]**	[1.081]**	[1.124]**
R^2	0.1	0.1	0.2	0.2	0.2
Dependent Variable: Δ ED6					
MP	0.265	0.292	0.339	0.35	0.345
	[0.175]	[0.180]	[0.133]**	[0.136]***	[0.138]**
Δ Score		1.526	3.389	4.032	3.91
		[0.952]	[1.898]*	[2.162]*	[2.289]*
R^2	0.0	0.1	0.1	0.1	0.1
Dependent Variable: Δ ED8					
MP	0.231	0.256	0.303	0.31	0.305
	[0.183]	[0.188]	[0.139]**	[0.147]**	[0.149]**
Δ Score		1.39	3.263	3.678	3.575
		[0.820]*	[1.665]**	[1.859]**	[1.966]*
R^2	0.0	0.1	0.1	0.2	0.1
Obs.	69	68	64	64	64
p-val. Q1=Q8		0.12	0.10	0.12	0.13

NOTES: *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 12: Regression Results for Univariate Model with Federal Funds Rates

	FF1	FF2	FF3	FF4	FF5
PANEL A					
RI_t	293.191	328.919	360.486	385.461	411.696
	[119.010]**	[108.990]***	[99.621]***	[89.889]***	[80.955]***
Obs.	68	68	67	66	66
logLik.	-448.34	-446.59	-438.52	-430.39	-428.77
PANEL B					
RI_t	18.099	23.305	35.784	47.835	68.156
	[5.342]***	[11.025]**	[17.417]**	[28.268]*	[36.198]*
RI_{t-1}	1.107	1.164	6.925	12.662	4.794
	[5.651]	[12.107]	[11.339]	[15.406]	[19.155]
MP_t	1.002	0.913	0.658	0.748	1.155
	[0.081]***	[0.110]***	[0.189]***	[0.310]**	[0.337]***
π_{t-1}	-0.005	0.002	-0.067	-0.071	-0.106
	[0.018]	[0.034]	[0.074]	[0.124]	[0.176]
Payroll_{t-1}	0.016	0.067	0.188	0.283	0.422
	[0.014]	[0.027]**	[0.059]***	[0.090]***	[0.125]***
f^n_{t-}	0.975	0.932	0.859	0.780	0.694
	[0.010]***	[0.022]***	[0.045]***	[0.072]***	[0.097]***
Obs.	67	67	66	65	65
logLik.	-240.74	-281.39	-296.8	-315.12	-328.33

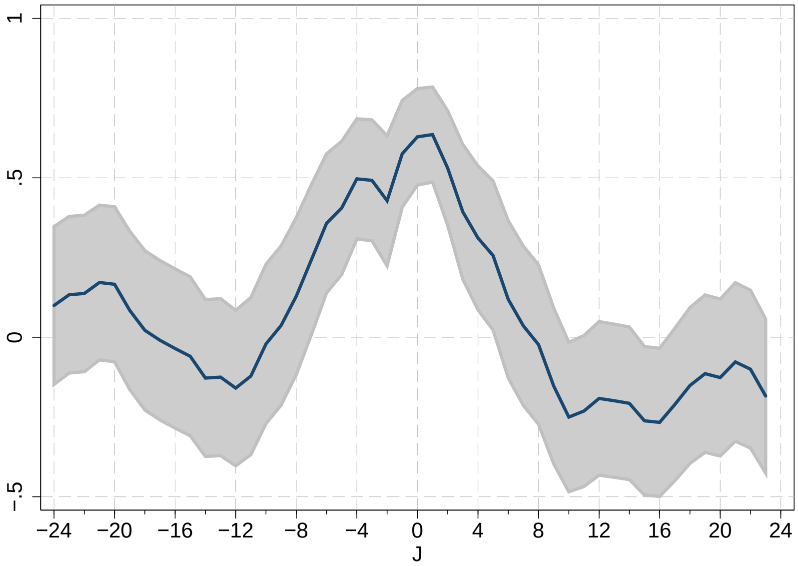
NOTES: *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 13: Regression Results for Univariate Model using Libor Rate

	ED1	ED2	ED3	ED4	ED5	ED6	ED7	ED8
PANEL A								
RI_t	371.006	433.392	456.203	412.303	313.376	264.066	197.168	148.239
	[97.983]***	[76.599]***	[60.598]***	[65.534]***	[101.922]***	[126.987]**	[162.169]	[180.877]
Obs.	69	69	67	65	64	61	59	57
logLik.	-452.18	-448.3	-434.03	-423.05	-418.45	-398.08	-384.07	-371.39
PANEL B								
RI_t	39.761	63.517	115.013	120.845	41.843	61.718	42.763	59.360
	[23.873]*	[45.563]	[56.640]**	[72.401]*	[73.839]	[78.638]	[68.160]	[72.250]
RI_{t-1}	-9.518	7.643	-28.143	-57.641	-34.873	-40.544	-9.310	-7.092
	[10.582]	[24.721]	[44.793]	[55.305]	[59.226]	[58.848]	[53.195]	[66.785]
MP_t	1.050	0.888	2.103	2.173	2.110	1.763	1.510	1.533
	[0.187]***	[0.434]**	[0.527]***	[0.565]***	[0.719]***	[0.760]**	[0.659]**	[0.642]**
π_{t-1}	-0.033	-0.141	-0.057	-0.149	0.035	0.170	0.311	0.540
	[0.062]	[0.198]	[0.342]	[0.519]	[0.613]	[0.706]	[0.564]	[0.471]
Payroll_{t-1}	0.141	0.575	0.954	1.326	1.442	1.374	1.211	0.943
	[0.033]***	[0.122]***	[0.205]***	[0.331]***	[0.404]***	[0.445]***	[0.356]***	[0.280]***
f_{t-}^n	0.900	0.620	0.310	-0.037	-0.334	-0.620	-0.908	-1.085
	[0.035]***	[0.102]***	[0.158]**	[0.217]	[0.254]	[0.273]**	[0.265]***	[0.228]***
Obs.	68	68	66	64	63	60	58	56
logLik.	-310.79	-359.95	-375.04	-380.51	-378.79	-361.16	-340.37	-325.81

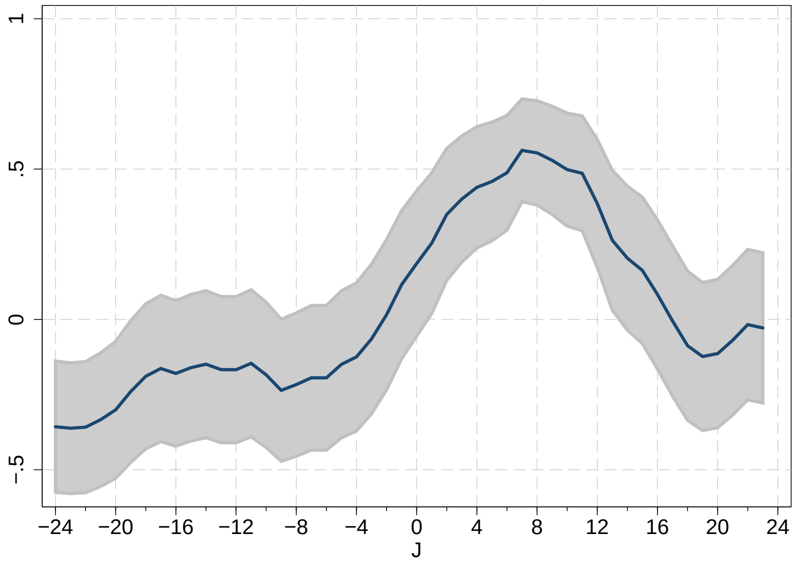
NOTES: *** significant at 1%, ** significant at 5%, * significant at 10%.

Figure 9: Cross-Correlation between the Taylor-rule residual and RI_{t-J}



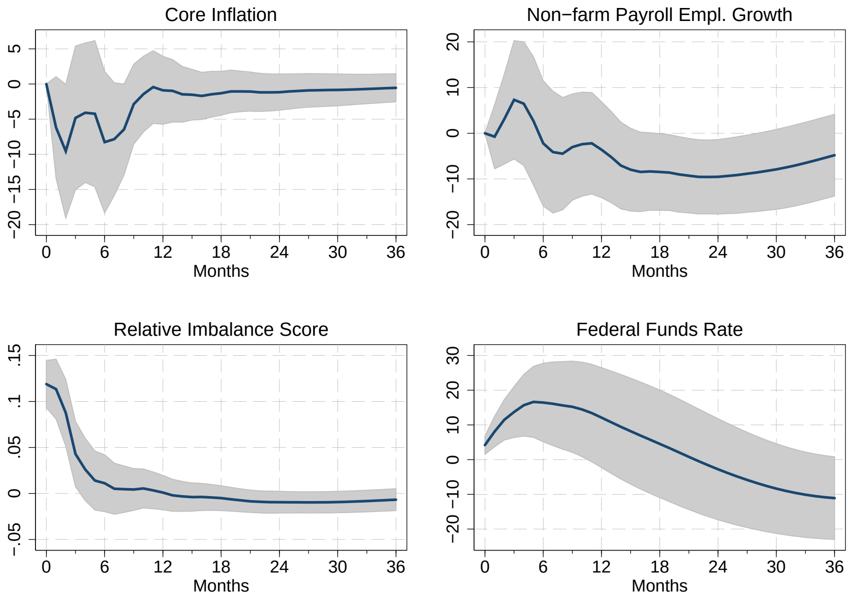
NOTES: The solid line denotes the correlation coefficient between the Taylor-rule residual, defined in Section 6.2, and RI_{t-J}, with the unit interval defined at an intermeeting frequency. Shaded areas represent two asymptotic standard-error confidence bands around the estimated correlation at each lead/lag.

Figure 10: Cross-Correlation between the Taylor-rule rate and RI_{t-J}



NOTES: The solid line denotes the correlation coefficient between the Taylor-rule rate, defined in Section 6.2, and RI_{t-J}, with the unit interval defined at an intermeeting frequency. Shaded areas represent two asymptotic standard-error confidence bands around the estimated correlation at each lead/lag.

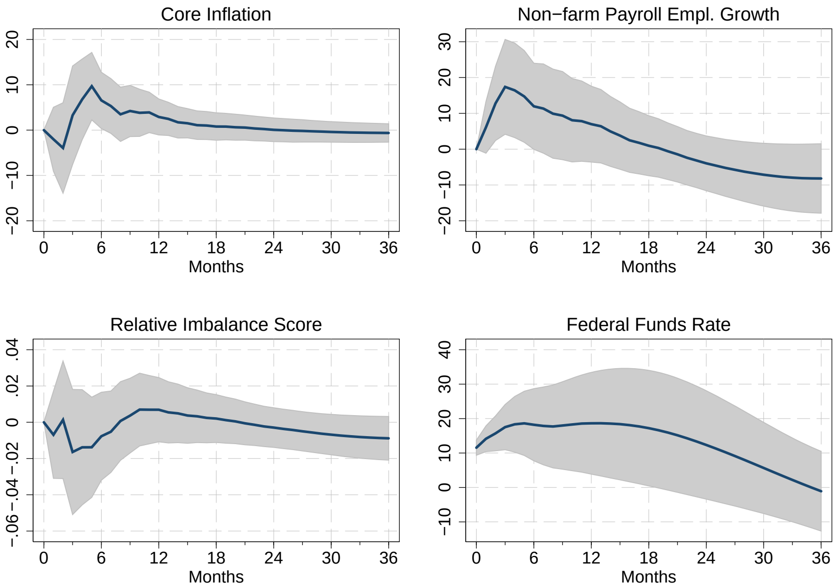
Figure 11: Impulse Responses of Core Variables to a Relative Imbalance Score Shock



NOTES: Shaded areas denote two-standard error bootstrapped confidence bands.

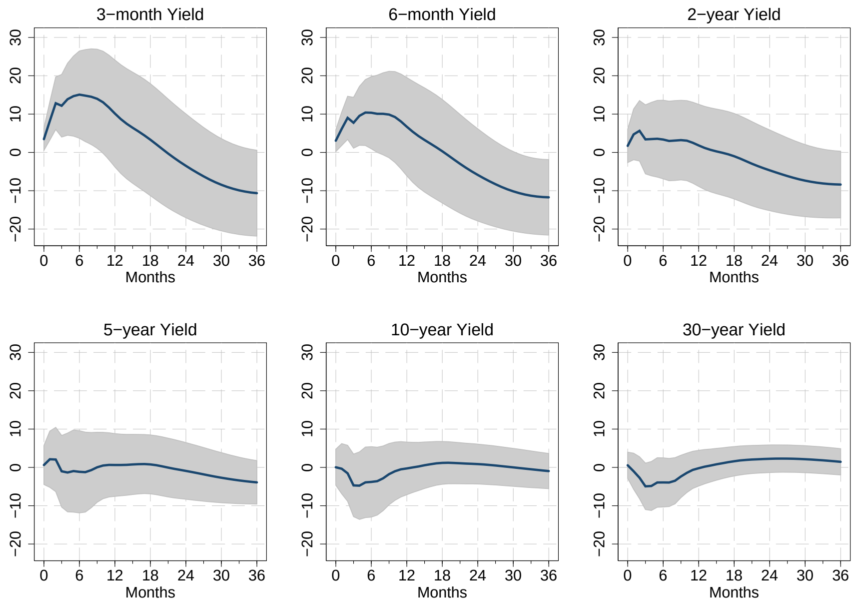
Figure 12:

Impulse Responses of Core Variables to a Federal Funds Rate Shock



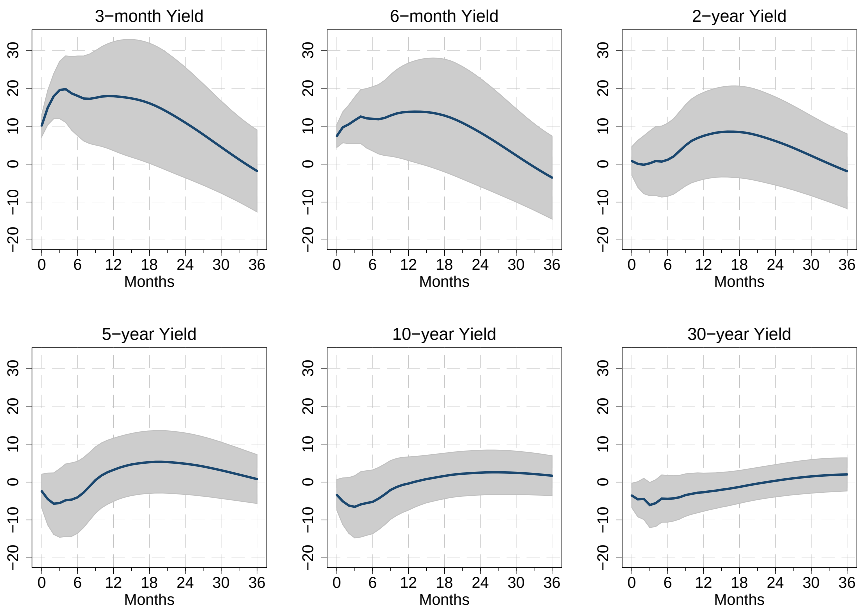
NOTES: Shaded areas denote two-standard error bootstrapped confidence bands.

Figure 13: Impulse Responses of Yields to a Relative Imbalance Score Shock



NOTES: Shaded areas denote two-standard error bootstrapped confidence bands.

Figure 14: Impulse Responses of Yields to a Federal Funds Rate Shock



NOTES: Shaded areas denote two-standard error bootstrapped confidence bands.

Table 14: Forecast Error Variance Decomposition for Core Variables

Variable	Inflation		Employment		RI		FFR	
Shock	RI	FFR	RI	FFR	RI	FFR	RI	FFR
Steps Ahead								
3-months	4.3	0.6	0.2	4.4	92.6	0.1	24.7	66.5
	[5.0]	[3.0]	[1.8]	[4.5]	[7.0]	[1.7]	[11.5]	[10.8]
6-months	5.6	5.1	1.7	15.1	87.6	1.6	30.5	51.7
	[6.2]	[4.9]	[5.5]	[10.1]	[9.1]	[3.9]	[14.6]	[13.7]
1-year	9.8	8	1.7	15.8	79.2	1.9	26	39.6
	[6.7]	[5.4]	[5.3]	[10.9]	[10.3]	[4.2]	[16.4]	[15.4]
2-years	10.1	8.5	7.9	14.3	75.4	2.1	13	31.8
	[6.8]	[5.5]	[7.5]	[10.1]	[11.0]	[4.5]	[14.3]	[15.8]
3-years	10.3	8.5	12	15.8	74.2	2.9	13.4	28.4
	[6.9]	[5.5]	[8.1]	[10.6]	[11.6]	[4.9]	[13.8]	[15.2]

NOTES: Numbers are expressed in percentage points. Bootstrapped standard errors reported in brackets.

Table 15: Forecast Error Variance Decomposition for Yields

Yield Maturity	3-month		6-month		2-year		5-year		10-year		30-year	
Shock	RI	FFR	RI	FFR	RI	FFR	RI	FFR	RI	FFR	RI	FFR
Steps Ahead												
	20.9	55.1	15	30.3	3.4	0	0.5	3.2	0.2	5.3	1.1	6.6
3-months	[10.8]	[10.3]	[9.6]	[9.7]	[6.4]	[3.0]	[4.3]	[5.4]	[3.6]	[6.4]	[3.9]	[7.4]
	24.3	53.6	17.6	31.5	3	0.1	0.4	4.2	2.6	7.7	5.2	10
6-months	[13.4]	[13.2]	[12.3]	[12.3]	[8.3]	[4.6]	[5.7]	[7.5]	[7.0]	[9.0]	[7.6]	[9.8]
	22.1	40.9	14.6	25.4	2.4	2.1	0.4	3.6	3.2	7.7	6.4	11.5
1-year	[15.2]	[14.9]	[13.2]	[13.4]	[8.8]	[8.0]	[7.1]	[7.3]	[8.7]	[9.4]	[9.1]	[10.6]
	11.4	32.7	6.8	21.5	1.7	7.3	0.3	6.6	2.9	7.3	6.8	11.5
2-years	[13.3]	[15.5]	[10.8]	[13.9]	[8.1]	[11.0]	[7.5]	[8.9]	[8.9]	[9.4]	[8.8]	[10.3]
	12.7	29.3	11	18.6	5.4	7.1	1.3	7.4	2.7	8.2	8.4	11.6
3-years	[13.0]	[15.0]	[10.8]	[13.1]	[8.5]	[10.9]	[7.4]	[9.2]	[8.7]	[9.7]	[8.7]	[10.5]

NOTES: Numbers are expressed in percentage points. Bootstrapped standard errors reported in brackets.