Federal Reserve Bank of Cleveland Working Paper Series

The FOMC versus the Staff: Do Policymakers Add Value in Their Tales?

Ilias Filippou, James Mitchell, and My T. Nguyen

Working Paper No. 23-20

August 2023

Suggested citation: Filippou, Ilias, James Mitchell, and My T. Nguyen. 2023. "The FOMC versus the Staff: Do Policymakers Add Value in Their Tales?" Working Paper No. 23-20. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202320.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

The FOMC versus the Staff: Do Policymakers Add Value in Their Tales?*

Ilias Filippou James Mitchell My T. Nguyen

This version: August 28, 2023

Abstract

Using close to 40 years of textual data from FOMC transcripts and the Federal Reserve staff's Greenbook/Tealbook, we extend Romer and Romer (2008) to test if the FOMC adds information relative to its staff forecasts not via its own quantitative forecasts but via its words. We use methods from natural language processing to extract from both types of document text-based forecasts that capture attentiveness to and sentiment about the macroeconomy. We test whether these text-based forecasts provide value-added in explaining the distribution of outcomes for GDP growth, the unemployment rate, and inflation. We find that FOMC tales about macroeconomic risks do add value in the tails, especially for GDP growth and the unemployment rate. For inflation, we find value-added in both FOMC point forecasts and narrative, once we extract from the text a broader set of measures of macroeconomic sentiment and risk attentiveness.

Keywords: Monetary Policy; Sentiment; Uncertainty; Risk; Forecast Evaluation; FOMC Meetings; Textual Analysis; Machine Learning; Quantile Regression.

JEL Classification: E17, E31, E52, E58.

***Acknowledgements**: Filippou (IliasFilippou@wustl.edu), Olin Business School, Washington University in St. Louis; Mitchell (James.Mitchell@clev.frb.org), Federal Reserve Bank of Cleveland; Nguyen (myn@wustl.edu), Olin Business School, Washington University in St. Louis. We thank Todd Clark, Ed Knotek, Pierlauro Lopez, and participants in a brownbag seminar at the Federal Reserve Bank of Cleveland for helpful comments. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

I. Introduction

Given its long and variable lags, monetary policy is forward looking and relies on forecasts of the future path of the macroeconomy. But whose forecasts should carry the most weight, and how and when are they informative? Romer and Romer (2008) famously questioned the value of the forecasts from the Federal Open Market Committee (FOMC) in the US. They concluded that FOMC forecasts do not contain useful information for forecasting GDP growth, inflation, and the unemployment rate relative to the FOMC's own staff's forecasts, as provided to the Committee ahead of each monetary policy meeting by the Board of Governors. Given these results, Romer and Romer provocatively suggested that FOMC members should focus on deciding between different policy options, taking as given (from their staff) the forecasts on which those policy options are based.

We revisit the question of how much value is contained in the FOMC's (the policymaker's, P) forecasts relative to the staff's (S), but we do so by extending the coverage of Romer and Romer (2008) to consider not only quantitative forecasts but also qualitative assessments of the economy. Text-based measures that capture attentiveness to and sentiment about the macroeconomy and its risks are extracted from the narratives provided in real time by both the FOMC and the staff alongside their respective quantitative forecasts. A growing literature has acknowledged the power of narratives in shaping economic outcomes (see Shiller (2017)). The considerable attention paid by both the FOMC, in its economic "go-around" during its meetings, and the staff, in the Greenbook/Tealbook narrative that accompanies their quantitative forecasts, suggests that there is perceived value in the forecast narrative. We empirically test the value-added of these words.

Our paper thus contributes to the now-considerable literature that uses textual-based measures to address various questions in macroeconomics. A notable strand of this literature, and one on which we draw, uses textual data to measure economic and policy uncertainty. Such measures are now widely used as time-varying measures of macroeconomic and economic policy uncertainty; for example, see Baker, Bloom, and Davis (2016). These measures typically involve counting the number of times *uncertainty*, and related words, appear in a body of text and then dividing by the total number of words in the document. More sophisticated model-based approaches, including deep learning models, have also been used to measure textual sentiment and quantify text (for example, see Hansen, McMahon, and Prat (2018) and Kalamara et al. (2022)). We too make use of a leading recent deep learning model developed by Google – bidirectional encoder representations from transformers (BERT) – pre-trained on millions of textual data points from Wikipedia and designed to better understand the meaning of words by using surrounding text to establish context and group similar words into topics. BERT is finding increasing use in economics (for example, see Gorodnichenko, Pham, and Talavera (2023)) given its high accuracy in interpreting texts.

Specifically, we use a range of textual methods from natural language processing (NLP) to extract and construct various measures of attentiveness to and sentiment about the macroeconomy from over 40 years of transcripts and narrative both from FOMC members themselves and from the Greenbook (the Tealbook since 2010).[1] The FOMC transcripts lend themselves to providing measures of risk and uncertainty, given the deliberative nature of the FOMC meetings. As Fed Chair Ben Bernanke quipped in 2015, "Monetary policy is 98% talk and 2% action."[2] At FOMC meetings, members discuss uncertainties about the data, uncertainties about the structural workings of the macroeconomy, the effectiveness of policy actions and current and future risks, and uncertainties facing the economy. This narrative may capture information not reflected in the point forecast that is nevertheless helpful when forecasting the distribution of possible outcomes. For example, FOMC members may discuss that the structure of the economy is changing and thus weakening confidence in their point forecasts or indicate that they are especially alert to macroeconomic risks either on the upside or the downside, information again not directly evident from their point forecasts alone.

[1]For ease, throughout this paper we refer to the publication (and its forecasts) as the Greenbook, even when it relates to the post-2010 period.

[2]See https://www.brookings.edu/blog/ben-bernanke/2015/03/30/inaugurating-a-new-blog/.

We compare the FOMC transcripts with the textual discussion that accompanies the projections of various economic indicators made by Board of Governors staff in the Greenbook, prepared as a briefing document about a week before each FOMC meeting. While statements from central banks have always been closely inspected by "Fed-watchers," increasingly methods from NLP have been applied to central bank text, both to enable and to automate the identification of topics that, for example, are informative about: policy preferences (for example, see Hansen et al. (2018), Peek, Rosengren, and Tootell (2016), and Shapiro and Wilson (2022)); central bank communication and financial market effects (Hansen, McMahon, and Tong (2019) and Gardner, Scotti, and Vega (2022)); macroeconomic forecasts (Balke, Fulmer, and Zhang (2017), Clements and Reade (2020), Lima, Godeiro, and Mohsin (2021), and Stekler and Symington (2016)); and the identification of monetary policy shocks and surprises (see Aruoba and Drechsel (2022) and Schmanski et al. (2023)). Another related paper is Sharpe, Sinha, and Hollrah (2023), who construct a single sentiment index from the Greenbook text and assess if it adds value relative to the Greenbook point forecasts. Our point of departure – in the spirit of Romer and Romer (2008) – is to assess the ability of a range of text-based methods to differentiate the information content of FOMC and staff narratives.

We then use textual data from the FOMC and the Board of Governors staff to revisit and update the regressions used in Romer and Romer (2008). Romer and Romer (2008) tested if the FOMC forecasts (P) contain useful information relative to the staff forecasts (S) by regressing the realized values of GDP growth, unemployment, and inflation on the two sets of point forecast (P and S): they estimate so-called Mincer and Zarnowitz (1969) (MZ) regressions. Rather than compare only the point forecast accuracy of the FOMC and Greenbook quantitative forecasts, we test whether the forecast narratives from both the FOMC and the Greenbook (what we call P* and S*) add value when forecasting the distribution of outcomes for GDP growth, the unemployment rate, and inflation. It is important to model the entire conditional distribution of the outcomes with respect to both the quantitative and the qualitative forecasts to allow for non-Gaussian features, to allow for the point forecasts to not necessarily be designed as conditional mean forecasts, and to reflect the possibility, for example, that FOMC and/or Greenbook discussions (say, of topics such as "risk") may better explain realizations in the tails of the distribution rather than the mean as captured by the linear MZ regressions estimated by OLS in Romer and Romer (2008). Segal, Shaliastovich, and Yaron (2015) and Adrian, Boyarchenko, and Giannone (2019) emphasize asymmetries in macroeconomic risks, between the left and right tails. Accordingly, we test if the informational content of the point and narrative forecasts from the FOMC and its staff varies by quantile of the macroeconomic outcome distribution.

Ours is not the first paper to revisit the Romer and Romer (2008) critique of FOMC (point) forecasts. Ellison and Sargent (2012) constructed their defense of the FOMC by developing a model whereby the FOMC is assumed to adopt a risk-management perspective. In their model, FOMC point forecasts (P) should be viewed and evaluated as worst-case not conditional mean forecasts. Binder and Wetzel (2018) also found that FOMC point forecasts do add value when economic conditions are more unfavorable. We operationalize such a "tails" defense of the FOMC via our quantile-based MZ regressions. But we supplement it with the "tales" defense that the FOMC also adds important information via its narrative (P*). Overall, we find that both FOMC and Greenbook corpora include important information for unemployment, real GDP growth, and inflation. The results are stronger for the first two variables, since more of the narrative about inflation is, in effect, already priced into the point forecast(s).

Our paper thus emphasizes the contrasting information content of FOMC and staff narratives. Computing a text-based measure of distance to determine how close the FOMC and Greenbook corpora are to each other in terms of their meaning, we show that the similarity between the two corpora across different macroeconomic topics identified by BERT is low. But similarity has, in general, been increasing over time. Our results have implications for the identification of monetary policy shocks when extending the approach of Romer and Romer (2004), as proposed by Aruoba and Drechsel (2022), to capture the information in both the point (numerical) forecasts and the forecast narrative. Our results imply that it is important to analyze both FOMC and Greenbook corpora.

The plan for the remainder of the paper is as follows. Section II first details the quantitative forecast data from the FOMC and the staff (P and S). Second, it introduces the textual data associated with the forecast narrative. Section III explains how we consider various methods from NLP to process the textual data and extract measures of the attentiveness to and sentiment about the macroeconomy (P* and S*). Importantly, we consider both simple word count methods and more sophisticated deep learning text classification algorithms that seek to understand the contextual meaning of words as well as their probability. Section IV then revisits and extends Romer and Romer's (2008) MZ regressions, adding in the text-based forecasts – P* and S* – and tests the value-added of the forecast narrative in explaining the distribution of outcomes for the unemployment rate, GDP growth, and inflation. Section V concludes. An online appendix contains supplementary tables and figures, as referenced in the main paper.

II. Forecast Data from the FOMC and the Greenbook

A. Quantitative Forecast Data

Our quantitative forecast data for both the FOMC (P) and the staff (S) are sourced identically to Romer and Romer (2008). As in their paper, the forecast data start in 1979 but the sample is updated from 2001, when the sample ends in Romer and Romer (2008), to 2017. Given the five-year embargo on publication, 2017 is the most recent year for which forecast data from P and S are currently available. We also follow Romer and Romer (2008) and define the outcomes or realizations data, used to measure the errors associated with the P and S forecasts, using real-time data released about three months after the end of the quarter of interest. These outturn data are typically taken from the Greenbook for the meeting following the data release.

Since 1979, the FOMC has prepared forecasts for inflation, unemployment, and real growth and published them in the Monetary Policy Report (MPR) that is submitted to Congress in February/March and June/July of each year. This is typically one or two weeks after the latest

FOMC meeting.[3] The forecasts made in February or March are for inflation and growth over the four quarters ending in the fourth quarter of the current year, and for unemployment in the fourth quarter of the current year. The forecasts made in June or July supplement these current-year forecasts with forecasts for the next year. The definitions of these variables have changed over time. Until July 1988, the inflation forecast was for the GNP implicit price deflator. Until July 1999, it was then for CPI inflation. Thereafter, it is for PCE inflation, until July 2004 when it switched to core PCE inflation (inflation excluding food and energy). For growth, real GNP was the target through July 1991, and thereafter, it was real GDP. We adapt our realizations data, as appropriate, to define the forecast error against the appropriate definition.

While each FOMC member submits his or her forecast, the MPR (and SEP) presents each member's forecasts only as a range and as a central tendency.[4] The range shows the highest and lowest forecasts of the individual members. The central tendency shows the highest and lowest forecasts after removing outliers, usually the three lowest and three highest forecasts. We follow Romer and Romer (2008) and use the midpoint of the central tendency. When this is not available, we use the midpoint of the range.[5]

Staff (S) forecasts for the same three variables are taken from the Greenbook, from 2010 renamed and repackaged as the Tealbook. These forecasts and the associated forecast narratives are prepared about a week *before* each FOMC meeting. The Federal Reserve does not explain how the Greenbook forecasts are produced, but they are believed to involve the use of both econometric models and judgment. FOMC members, therefore, have a timing advantage, since they can, if they wish, in effect condition their P forecasts on those from the Greenbook (S):

[3] These forecasts are available at https://www.federalreserve.gov/monetarypolicy/publications/mpr_default.htm. Since October 2007, the FOMC has released forecasts associated with four FOMC meetings per year in the Summary of Economic Projections (SEP). In this paper, we analyze the bianannual MPR data, as these are available over the longer sample back to 1979. Since the inception of the SEP, the forecasts in the MPR are those from the latest SEP.

[4] That is, individual FOMC forecasts are not published in the MPR. Individual forecasts, however, are now available over a shorter sample from 1992 and with a 10-year release delay restricting public access to more recent forecasts; see Romer (2010) and Banternghansa and McCracken (2009). FOMC members make their forecasts conditional on their own judgment of "appropriate monetary policy."

[5] An extension is to explore entering the forecast range/interval from the FOMC into the MZ regressions to test if this measure of "disagreement" adds value.

FOMC members know the staff forecasts before they submit theirs. The Greenbook forecasts are commonly believed to be modal forecasts rather than conditional mean forecasts. We return in Section IV to the importance of evaluating forecasts under alternatives to quadratic loss.

Previous work evaluating the Greenbook forecasts has found that, in general, they are of good quality and outperform econometric model and private-sector forecasts, if not consistently over time then certainly for some sample periods; for example, see Romer and Romer (2000), Sims (2002), and Faust and Wright (2009). Notwithstanding Romer and Romer's (2008) critique of the FOMC's forecasts – and the focus of this paper – the FOMC's forecasts have also been separately analyzed. Arai (2016) finds the FOMC inflation forecasts to be better than those for GDP or unemployment.

B. Textual Forecast Data: The Forecast Narrative

The FOMC typically meets eight times a year. To match the Romer and Romer (2008) sample, and relate the textual data to the corresponding quantitative forecast (discussed in Section II.A), we focus on those FOMC meetings and Greenbooks most closely timed with the biannual quantitative forecasts that the FOMC provides to Congress. Since publication of the SEP started in 2007, this means that the FOMC quantitative forecasts in the Febuary/March MPR are a little old, as they are from the December SEP. The mismatch for the June/July MPR is smaller, since those forecasts are associated with the June SEP.

The deliberations of the FOMC were conducted in secret for decades. Not until March 1994 did the Federal Reserve release historical transcripts based on tape recordings of meetings. Henceforth transcripts were, and still are, released with a five-year lag.[6] We focus our textual analysis on the transcripts rather than on other FOMC documents, since the transcripts contain the most detailed information about FOMC discussions. We note that the FOMC text we examine includes little or no quantitative information (we do not include any tables or figures

[6]See https://www.federalreserve.gov/monetarypolicy/fomc/historical.htm for the source of the FOMC transcripts.

in our textual analysis).

As well as presenting to the FOMC the quantitative forecasts of the staff of the Board of Governors, the Greenbook provides the staff's forecast narrative. The length and content of the Greenbook have evolved over time. Over our sample, the Greenbook has been divided into two parts. We focus on "Part 1," which summarizes economic developments in the US and abroad, since this is where the staff's forecasts are presented (although we drop any tables or figures from the subsequent textual analysis). The length of Part 1 grew over time and was around 50 pages long by 2010 when the Tealbook replaced the Greenbook. From 2010, when we use the Tealbook, we focus on "Tealbook A," since this is again where the staff provide their in-depth narrative as well as their quantitative forecasts. Tealbook A is a longer document, typically around 100 pages, since it covers both Parts 1 and 2 of what was the Greenbook. There is no obvious way to divide Tealbook A into the two parts that previously comprised the Greenbook; hence our textual methods are applied to all of the text in Tealbook A.

Our textual analysis focuses on 77 FOMC transcripts and 77 Greenbooks starting from July 1979 through June 2017. Over this sample period, the number of words in the FOMC transcripts increased from 5,401 to 15,531. The number of words in the Greenbook also increased, from 2,244 to 12,371. The average (over the 77 documents) number of words in the FOMC transcripts and the Greenbooks is 12,840 and 4,698, respectively.

III. Construction of Text-Based Factors

This section sets out the textual methods used to measure attentiveness to and sentiment about the economy. As discussed, we break these measures into those of attentiveness and those of sentiment. We emphasize that all of these measures can be computed in real time. By pre-selecting the words (oriented around risk and uncertainty) we use, we minimize text-based "information leakage." This occurs when one benefits from look-ahead bias when deciding what words are informative, for example, by using the entire corpus (over time) to identify the most popular words. Often it is the case that such words would not have been identifiable in real time. The only exception is that for BERT, we do use the full (pooled over time) sample of transcripts or Greenbooks (over time) to identify both "risk" and macroeconomic topics related to unemployment, GDP growth, and inflation. But then, as explained below, we do estimate the frequency and tone of these topics at a given point in time. In other words, we use BERT to identify the paragraphs, in a given corpus, that discuss the different topics.[7]

Specifically, we estimate model-free and model-based measures of attention and sentiment from the FOMC transcripts and the Greenbook. Our main focus is on text-based measures of "risk," given our prior expectation that the narrative about risk (and related subjects) is an important way that each corpus communicates risks and uncertainties about the macroeconomy. To this end, we construct n-grams of risk (model-free) and use BERT to identify risk topics for both corpora.

We also build model-free and model-based measures of sentiment. The model-free methods use dictionaries from Loughran and McDonald (2011) and Sharpe et al. (2023). Model-based estimation of the tone of each document uses BERT.

A. Pre-Processing: Data Cleaning

Before transforming the text into numbers, we must pre-process the text. This data cleaning proceeds as follows.

Stop words. We eliminate words that provide very little information, and we label them as stop words. For example, words such as "I," "the," "and," "a," "she," and "he" are included in the set of words that are removed from our analysis. In addition, we eliminate words that are very rare or that appear very frequently in the transcripts. We also only consider words with at least four letters.

[7]For example, the risk sentiment measure from BERT is based on the sentiment of the paragraphs that discuss risk topics as classified by the BERT model.

Linguistic stemmer. We also tokenize our corpus by applying Porter's (1980) stemmer, which eliminates the suffixes (for example, "-tion") of every word in our sample. In other words, our bag of words contains only the root words included in the transcripts in order to avoid the same word appearing twice in our sample.

Other filters. We eliminate all numbers, punctuation, and special characters and convert all strings to lowercase.

B. N-grams of Risk Attentiveness

Our first set of textual analysis methods is based on a common word categorization or bag-of-words approach used to measure the time-varying attention given to "risk-" related topics. In this method, every document is described by a vector of word counts that construct a term-document matrix. Specifically, the document-term matrix is defined as the frequency of terms that occur in a group of documents where rows correspond to documents in the group and columns include the unique terms of the documents. By tracking the words of interest ("risk-" related ones in our case), we estimate the time-varying attention given in the corpus to a specific topic.

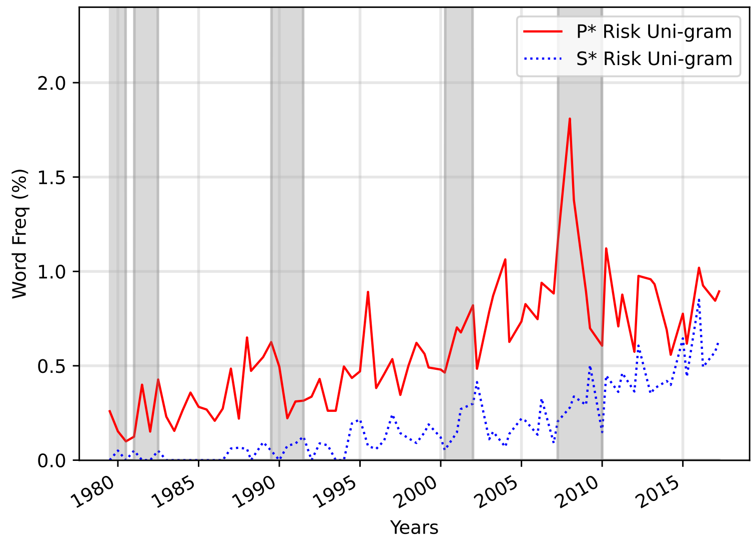
Uni-grams: Risk

In order to measure the level of risk expressed by the staff and the policymakers, we first construct a risk measure in the spirit of Baker et al. (2016) and Hansen and McMahon (2016) by taking a simple count of the word "risk" in each document. The measure is normalized by the total number of words in each document.[8] Specifically:

$$\text{Risk}_t = \frac{n_t^{\text{Risk}}}{n_t^{\text{Doc}}} \tag{1}$$

[8]Our risk measures do not attempt to distinguish risk about the *current* economic climate from risks specific to the forecast (the *future*, as opposed to the present or the past).

Figure 1. Uni-grams: Risk



The figure displays the frequency of risk uni-grams, estimated via (1). We show results for the FOMC transcripts (or P*) (red) and the Greenbook (or S*) (blue). Shaded areas represent NBER recessions. The data are semiannual from 1979 to 2017.

where n_t^{Risk} represents the total number of times that the word "risk" is mentioned in the corpus and n_t^{Doc} is the total number of words in the time t corpus. We denote the risk uni-grams that are constructed based on the FOMC (Greenbook) corpus by P Risk$_t$ (S Risk$_t$). Figure 1 shows the frequency over time of the word "risk," (1), in the FOMC (or P*) and the Greenbook (or S*), respectively. In line with our conjecture, we find that the FOMC commonly mentions the word risk in its discussions, and the staff also use this word extensively in the Greenbook, especially during periods of financial stress such as the global financial crisis of 2008-2009. We also observe that the frequency of the word risk increases over time and is higher in the FOMC transcripts than in the Greenbook.

Bi-grams and Tri-grams: Risk

We also consider the frequency of consecutive words that include the term "risk." We eliminate stop words and tokenize the corpora before performing this exercise in order to obtain more meaningful bi-grams and tri-grams. The choice of these bi-grams and tri-grams is based on the risk-related pairs or combinations of three words that appear more often in the whole (pooled across time) universe of corpora. We perform this exercise separately for the two corpora.

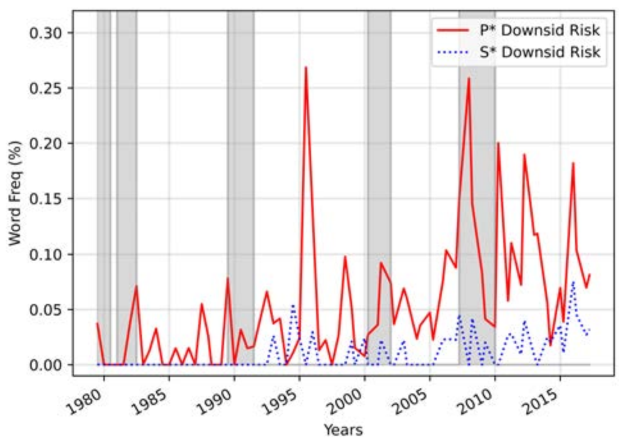
Bi-grams. We compute the frequency of the top 10 bi-grams of the words "risk" in the FOMC and Greenbook corpora. Specifically, we compute the number of times that top bi-grams, such as downside risk, policy uncertainty, and inflation risk, are mentioned in each corpus over time. The results are shown in Table A1 (in the online appendix), and we summarize the findings here. Interestingly, we find that the top bi-grams in the FOMC transcripts largely coincide with those in the Greenbook corpus. The most important bi-grams are related to "downside risk." We also find that "upside risk" is particularly important in both corpora. This is in line with Adrian et al. (2019), who highlight that upside risks to GDP growth tend to be lower in most periods, while downside risks become pronounced as financial conditions deteriorate. This is further verified in Figure 2, where the frequency of the bi-gram "downside risk" increases over time in comparison to "upside risk," which exhibits lower frequencies but remains important.

Tri-grams. Figure 3 reports the frequency of the top two tri-grams of the word "risk" for each corpus. Regarding the FOMC corpus, the top tri-grams are "upside risk inflation" and "downside risk growth." This perhaps reflects the increasing focus of the FOMC, for an inflation-targeting bank, on inflation movements and the state of the economy. The top two tri-grams in the Greenbook corpus are "equity risk premium" and "risk premium corporates"; this reflects the attention of the staff not only to macro topics but also to movements in the stock market. This is not surprising, as monitoring systematic risks is also a central policy objective.

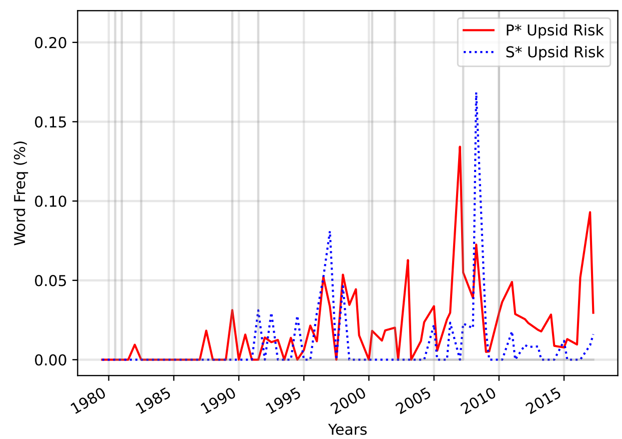
Risk Synonyms

We also construct an alternative measure of risk and uncertainty based on the dictionary of Hassan et al. (2019, HHLT). It is worth noting that this lexicon includes the words risk and

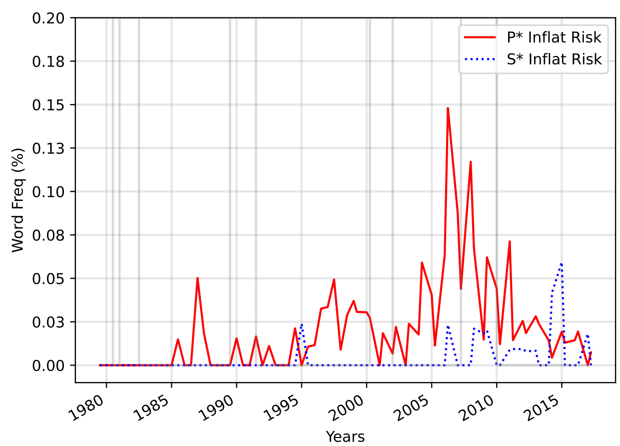
Figure 2. Bi-grams: Risk



(a) Downside Risk



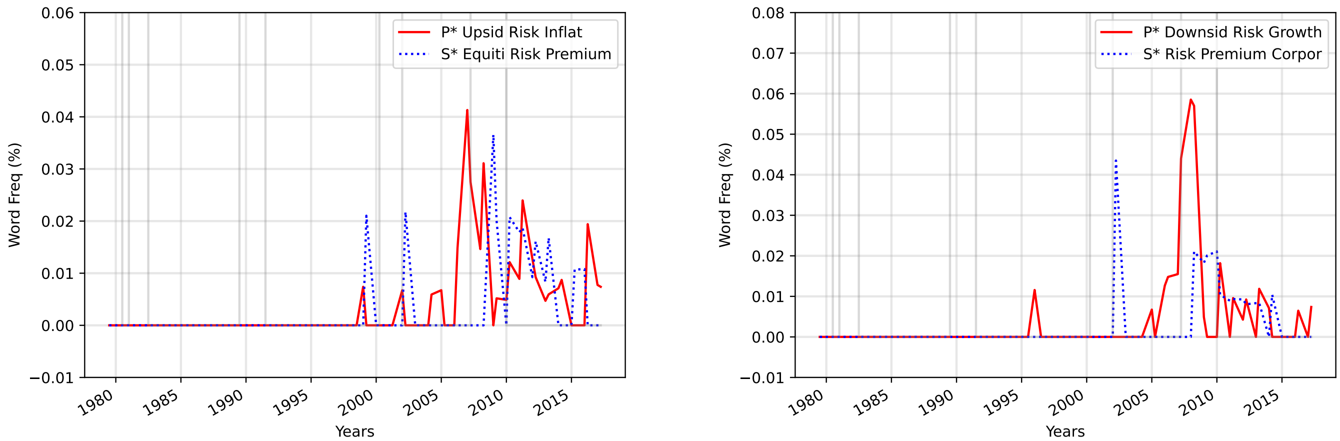
(b) Upside Risk



(c) Inflation Risk

The figure displays the top 3 bi-grams of the word "risk" in the FOMC and Greenbook corpora. We show results for FOMC (or P*) (red) and Greenbook (or S*) (blue) corpora. Shaded areas represent NBER recessions. The data are semiannual from 1979 to 2017.

uncertainty and their variants. Given its broader dictionary, this measure may behave differently from our risk uni-grams. It is more likely to exhibit similarities with the aforementioned measure when "risk" alone is heavily mentioned in the text. The HHLT risk measure is defined as:



(A) Top One Tri-Grams

(B) Top Two Tri-Grams

Figure 3. Tri-grams: Risk

The figure displays the top 2 tri-grams of the word "risk" in the FOMC and the Greenbook corpora. We show results for FOMC (or P*) (red) and Greenbook (or S*) (blue) corpora. Shaded areas represent NBER recessions. The data are semiannual from 1979 to 2017.

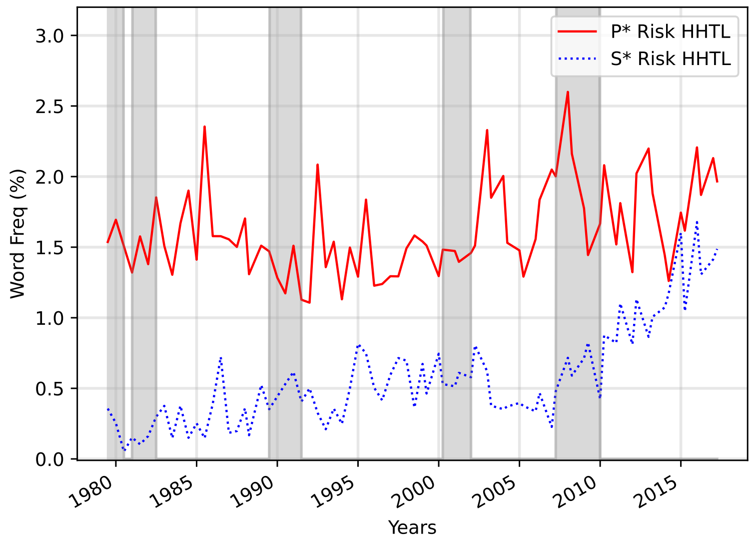
$$\text{Risk}_t^{\text{HHLT}} = \frac{n_t^{\text{Risk}^{\text{HHLT}}}}{n_t^{\text{Doc}}} \tag{2}$$

where $n_t^{\text{Risk}^{\text{HHLT}}}$ represents the total number of uncertainty words of the Hassan et al. (2019) dictionary and n_t^{Doc} is the total number of words of the transcript at time t.

We denote by P* Risk HHLT$_t$ (S* Risk HHLT$_t$) the risk estimates that are constructed based on words in the Hassan et al. (2019) dictionary that appear in the FOMC (Greenbook) corpus. Figure 4 displays the time-series of the risk measure for the two corpora. We find that both measures spike around financial crises (shaded areas denote NBER recessions) and that the FOMC more frequently uses words related to not just risk but uncertainty too.[9]

[9]It is worth noting that the upward trend to the Greenbook risk measure at the end of the sample may be partly mechanical, due to the introduction of a "risk and uncertainty" section starting with the June 2010 Greenbook.

Figure 4. HHTL Risk



The figure displays frequency of the Hassan et al. (2019) risk measure, estimated via (2). We show results for FOMC (or P*) (red) and Greenbook (or S*) (blue) transcripts. Shaded areas represent NBER recessions. The data are semiannual from 1979 to 2017.

C. BERT–Based Estimates of Risk Attentiveness

As an alternative to specifying a dictionary, we use BERT: a deep neural network topic-based model developed by Devlin et al. (2018). The motivation for BERT is that the NLP literature highlights the success of capturing the complex dynamics of language by focusing on sequences of words instead of examining words in isolation. BERT achieves this by representing words as embeddings, which are high-dimensional vector-space models of text where each unique word in a corpus is expressed as a vector in a shared vector space. This means that BERT can identify whether a number of texts have similar meanings, regardless of the common words they share. Another feature is that it can capture dependencies between words. For example, the word "downside" may depend on the word "risk."

The main variants of embeddings, such as Word2Vec (Mikolov et al. (2017)), GloVe (Pennington, Socher, and Manning (2014)), and FastText (Joulin et al. (2016)), use the distributional hypothesis to identify relationships in the embedding space (Harris (1954)). This hypothesis suggests that semantically similar words tend to have similar distributions and so appear in related linguistic contexts. However, these approaches have certain shortcomings. Specifically, they focus on word-level embeddings because they do not perform well as sentence encoders in the sense that they usually misinterpret context. Thus, they use as an input to the model one word, and the output is a vector representation of that word (Perone, Silveira, and Paula (2018)). BERT instead focuses on contextual embeddings, so the input to the model is a sentence instead of a single word. It is also directional, which means that it takes into consideration both preceding and subsequent context to generate the embeddings of a word, in contrast to unidirectional models (like ELMo and ULMFit). For this reason, BERT can interpret texts with more precision and has been increasingly employed in other studies; for example, see Chava, Du, and Malakar (2021) and Gorodnichenko et al. (2023).

Other unsupervised topic models are also finding growing applications in macroeconomics, including latent Dirichlet allocation (LDA) and latent semantic indexing (LSI). But, despite their widespread usage, LDA and LSI rely on the bag-of-words representation of documents, implying that word ordering and semantics are overlooked. They also are computationally intensive, given that they need to be trained each time they are used. This further explains our preference for BERT.

The BERT model is pre-trained, which implies that it works better out-of-the-box. Such "transfer learning" from previous applications of the model is important for the documents in earlier years of our corpus that are smaller.[10] In addition, the model is contextualized, meaning it builds a vector for each word based on its context. This is based on the idea that the utilization of a word (for example, syntax and semantics) depends on its context. For example, the word "play" has different meanings in the following sentences extracted from our corpus (to illustrate we use the transcripts for the FOMC meeting in April 2016): "Maybe there was agreement on overvaluation of housing, but exactly how that would *play* out and exactly how

[10] The neural network of BERT is pre-electronic trained on 800 million BooksCorpus and 2,500 million Wikipedia words. Thus, the model is able to identify words that have similar meanings based on pre-training.

different approaches to monetary policy would affect that were, I think, quite uncertain"; and "the strong role *played* by enhanced capital requirements in our current regulatory approaches is an excellent step forward. Regular stress testing is another strong tool." BERT aims to capture such differences in meaning. BERT reduces the number of unique words that are included in the model by partitioning each word into smaller tokens (such as subwords). It is also designed to encode entire sentences with a length of 512 tokens.

Our goal with BERT remains to extract the "risk" topic(s) from the FOMC transcripts and Greenbooks but to let the topic model decide which words to place in this topic based on their linguistic similarity. Inspection of the most frequent words in the topic then provides insight into the nature of the topic, and helps the researcher "identify" it.

We apply BERT separately to the FOMC and Greenbook corpora. We do so at the paragraph level, for a given (time t) corpus. After we have converted the text documents into embeddings, we run a clustering model on the embedded transcripts.[11] Our goal is to identify paragraphs with discussions about risk. To this end, having estimated BERT we identify risk topics from each corpus at time t. We report word clouds showing the most prominent words in the risk topic in Figure 5 using full sample ($t = 1, ..., T$) information. The left graph shows the most important words in the risk topic from the FOMC transcripts, and the right graph shows the most prominent words in the Greenbook risk topic. We find that the words with the highest probabilities in the two word clouds include words such as risk, uncertainty, downside risk, upside risk, inflation risk, tail, recession, recovery, weaker, turbulence, spillover, and damage. This highlights the ability of our BERT model to identify paragraphs that do indeed discuss risk-related subjects.

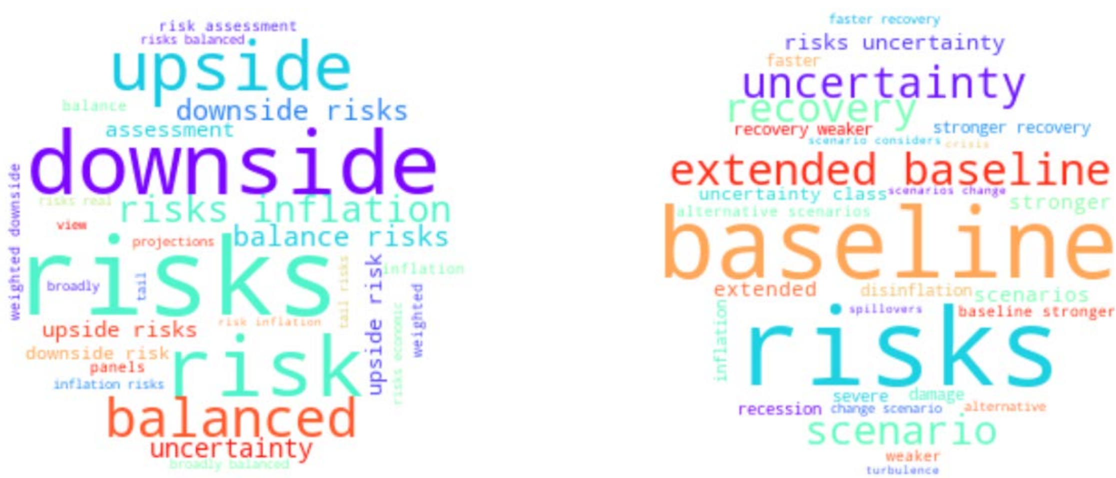
In order to have a measure of risk that is comparable to our n-grams, we then compute the frequency of paragraphs in the time t corpus with risk content. The BERT-based risk measure

[11] There are different variants of BERT embeddings based on the training data and the architecture. We consider word embeddings that are created from the BERT base model (12 layers, 768 hidden states, 12 heads, and 110 million parameters).

is accordingly defined as:

$$\text{Risk}_t^{\text{Bert}} = \frac{n_t^{\text{Risk}^{\text{Bert}}}}{n_t^{\text{Doc}}} \tag{3}$$

where $n_t^{\text{Risk}^{\text{Bert}}}$ represents the total number of paragraphs that belong to the "risk" topic with high probability and n_t^{Doc} is the total number of paragraphs in the corpus at time t. Figure 6 plots the risk estimates from BERT for both corpora. Again we see differences, not just between P* and S*, but between the estimates of P* and S* in this figure versus the previous figures. But Figure 6 does again show that it is the FOMC that appears to have been paying more attention to risk over time, with P* trending upward over time, albeit spiking during the post-2000 recessions.

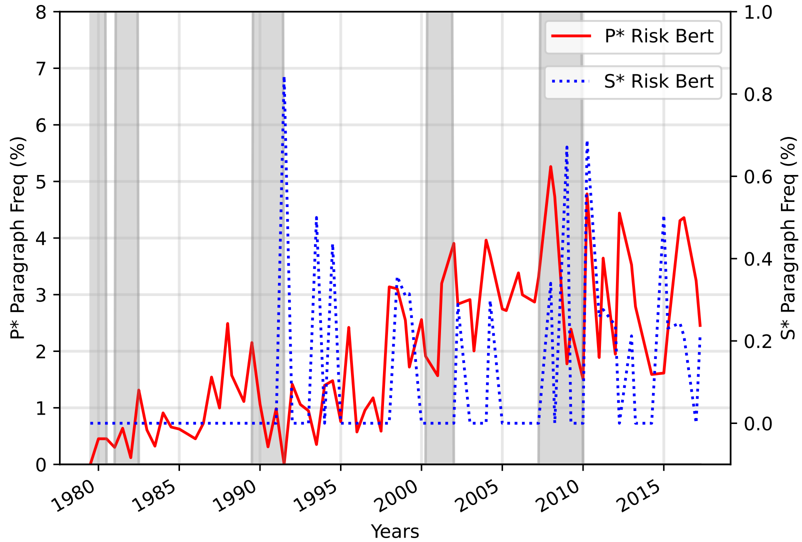


(A) FOMC (B) Greenbook

Figure 5. Risk Topic Word Clouds from BERT

The figure displays word clouds for the risk topic extracted from the FOMC transcripts and the Greenbook. We show results for the FOMC (or P*) (left panel) and Greenbook (or S*) (right panel) corpora. The data are semiannual from 1979 through 2017.

Figure 6. Uni-grams: Risk Estimates using BERT



The figure displays the frequency of paragraphs that discuss risk topics based on BERT, estimated via (3). We show results for the FOMC (or P*) (red) and Greenbook (or S*) (blue) corpora. Shaded areas represent NBER recessions. The data are semiannual from 1979 to 2017.

D. FOMC and Greenbook Tone about the Macroeconomy

To complement these text-based measures of risk attentiveness in Sections III.B and III.C, we now measure the tone or sentiment about the macroeconomy of each corpus. Again we consider model-free and model-based approaches.

Dictionary-Based Tone

We first measure the tone of the FOMC transcripts and the Greenbook in a model-free manner following a bag-of-words methodology as in Tetlock (2007) and Loughran and McDonald (2011). In particular, we calculate the tone of each document by computing the frequency of keywords that appear in a tone lexicon calibrated to financial data. Loughran and McDonald (2011) recognized that the negative words included in the widely used Harvard IV-4 Psychosociological Dictionary (for example, the Harvard-IV-4 TagNeg (H4N) file) might not reflect the tone of financial (or indeed macroeconomic) text. For this reason, the authors offer an alternative dictionary that is constructed based on 10-K filings and is able to capture the tone of documents with financial contexts. Thus, we measure the tone of each of our corpora as the difference between the number of positive and negative tonal words. Intuitively, a higher tone indicates a more positive or less negative sentiment from the FOMC or staff. The measure takes the form:

$$\text{Tone}_t^{\text{LM}} = \frac{n_t^{\text{Negative}} - n_t^{\text{Positive}}}{n_t^{\text{Doc}}} \tag{4}$$

where n_t^{Negative} represents the total number of negative words in the corpus at time t, and n_t^{Positive} represents the total number of positive words in the corpus at time t. The tone score is then defined as the difference between positive and negative word counts divided by the total number of words in each corpus. Figure 7 plots the tonal estimates using the LM dictionary. We see that, as expected, net negativity as expressed by the FOMC narrative clearly rises during the Great Recession, but not all previous recessions. The tonal estimates from the staff (S*) align with those from P* pretty well, but the relationship appears to weaken during the 2007-9 recession when S* does not spike like P*. Recall how both of these tonal measures seen in Figure 7 assess the sentiment of the entire document, not the tone of specific topics to which we now turn.[12]

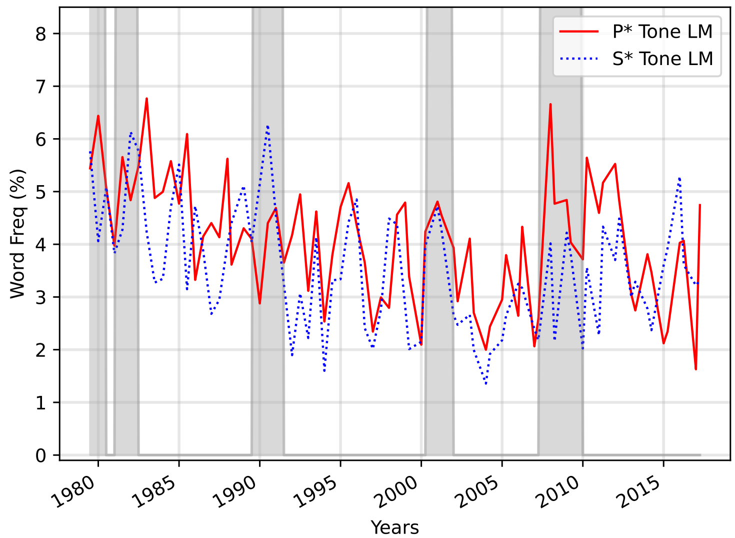
Model-Based Tone

The previous section used a dictionary-based approach to estimate the tone of each corpus. To minimize the subjective use of judgment in choosing the dictionary, here we construct an alternative model-based measure of tone. Specifically, we use FinBERT, which is a variant of BERT that allows us to capture the sentiment in a dynamic way (see Yang, Uy, and Huang (2020)). FinBERT is pre-trained on financial texts.[13] We use FinBERT instead of BERT –

[12]In the online appendix – see Figure A1 – we contrast the LM tone-based estimates seen in Figure 7 with those when we use the dictionary of Sharpe et al. (2023), who construct their own dictionary of positive and negative words. There are pronounced differences, reminding us once more – a theme of our paper – of the importance of consulting a variety of text-based methods given that a priori it is not clear which single measure should be preferred.

[13]The corpus used consists of the following documents: corporate annual and quarterly 10-K and 10-Q filings of Russell 3000 firms between 1994 and 2019; financial analyst reports issued for S&P 500 firms between 2003 and 2012 from the Thomson Investext database; and earnings conference call transcripts of 7,740 public firms

Figure 7. LM Tone Measure



The figure displays the Loughran and McDonald (2011) tone measure of the FOMC and Greenbook corpora, estimated via (4). We show results for the FOMC (or P*) (red) and Greenbook (or S*) (blue) transcripts. Shaded areas represent NBER recessions. The data are from 1979:Q4 to 2017:Q2.

which is pre-trained with general texts – to have a training sample more analogous to the corpus that Loughran and McDonald (2011) employ to generate their dictionary.

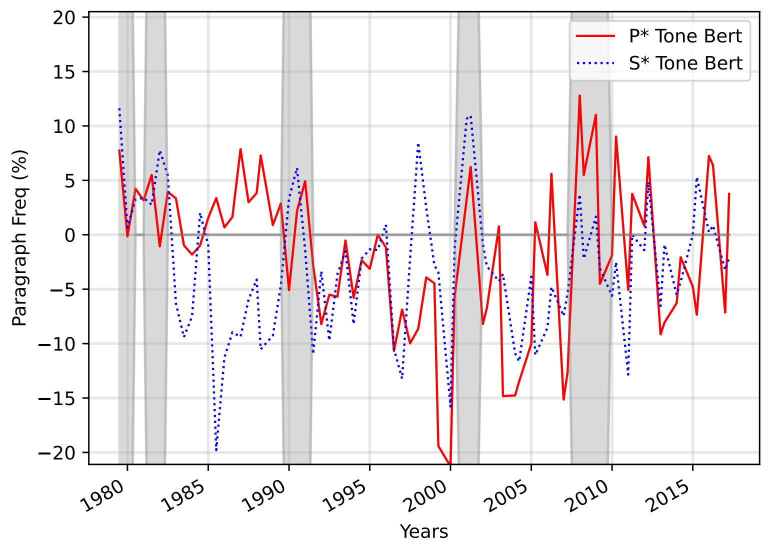
We apply FinBERT to the FOMC transcripts and Greenbook at the paragraph level. Specifically, at time t, we partition the corpus into paragraphs and apply FinBERT. We measure the tone of each corpus as the difference between the number of positive and negative tonal paragraphs at time t. Intuitively, a higher tone indicates more negative sentiment. Specifically, the measure takes the form:

$$\text{Tone}_t^{\text{FinBERT}} = \frac{n_t^{\text{Negative}} - n_t^{\text{Positive}}}{n_t^{\text{Doc}}} \tag{5}$$

where n_t^{Negative} represents the total number of negative paragraphs of the corpus (either the transcripts or the Greenbook) at time t, and n_t^{Positive} represents the total number of positive paragraphs of the corpus at time t. Tone score is defined as the difference between positive and negative paragraph counts, divided by the total number of paragraphs. It is worth noting that

between 2004 and 2019 from the SeekingAlpha website.

Figure 8. Tone Based on BERT



The figure displays the BERT-based tone index, estimated via (5). We show results for the FOMC (or P*) (red) and Greenbook (or S*) (blue) corpora. Shaded areas represent NBER recessions. The data are semiannual from 1979 to 2017.

the measure is estimated in real time so there is no look-ahead bias. We plot the estimates in Figure 8 and note how the BERT-based estimates of tone appear to exhibit greater concordance with NBER recessions than those using the LM dictionary seen in Figure 7.

Tone of Macro Topics Using BERT

We next use BERT to extract topics that are directly related to GDP growth, unemployment, and inflation. Specifically, we first apply BERT to extract topics associated with each of these three macroeconomic variables. Then we calculate the LM sentiment of the paragraphs that discuss each of these topics. These sentiment measures are, therefore, local to each topic, rather than based on the entire corpus that likely involves a mix of different economic concepts. Figure 9 plots the estimates.[14] We see that while sentiment about unemployment, GDP growth, and inflation often deteriorates in recessionary periods, there is considerable heterogeneity across

[14]In a different context, Filippou, Nguyen, and Viswanath-Natraj (2023) employ a similar methodology to extract the most important topics from cryptocurrency news.

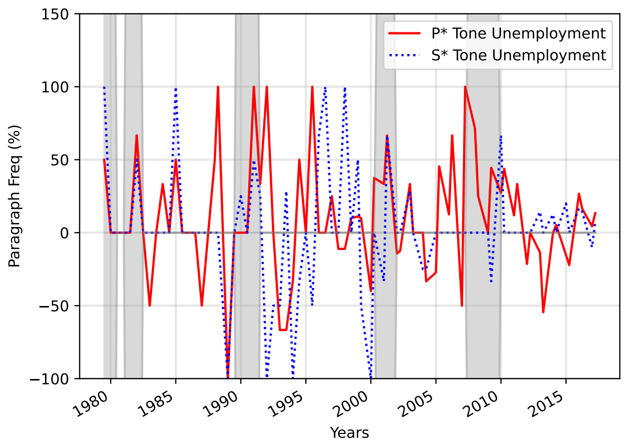
the sentiment measures in panels (A) through (C) of Figure 9. Table A2 in the online appendix, which reports correlation coefficients across the different textual measures *within* the FOMC transcripts and the Greenbook, confirms that the correlations between these three measures - in a given corpus - are low. Only tone GDP has a statistically significant correlation with tone inflation and tone unemployment in the FOMC transcripts.

E. Correlation between FOMC and Greenbook Textual Measures

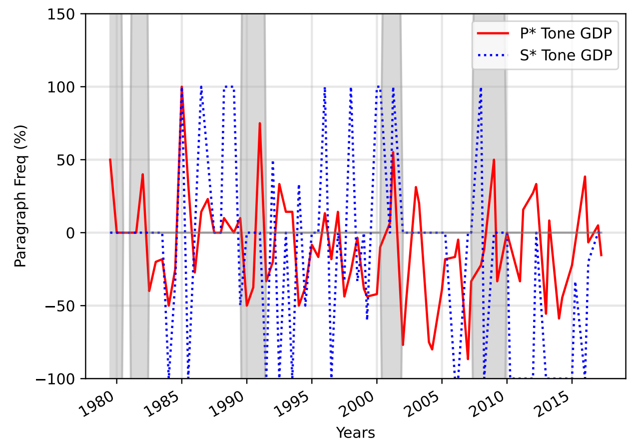
To summarize the relationship between each of these text-based factors across the FOMC and Greenbook corpora, Table 1 reports their correlation coefficients. The size of these coefficients confirms our graphical analysis: the textual measures from P* and S* are often quite distinct. Interestingly, Table 1 shows that all the factors are positively correlated across the FOMC and the Greenbook. However, the strength of the correlation varies by measure, indicating the potential semantic differences between the two corpora as well as the importance of differentiating between the FOMC and Greenbook documents.

Table 1 shows that the most correlated measure across the FOMC and the Greenbook, by quite a margin, is the risk uni-gram, (1). Specifically, these two risk measures exhibit a correlation of 0.61 which is statistically significant. Thus, there is strong commonality in the attention given to risk in the two corpora, but important differences remain. As illustrated in Figure 1, we see that the FOMC mentions the word "risk" more often and its measure tends to be more volatile. However, other risk-based measures are correlated much more weakly, suggesting that the narratives in the two corpora are quite distinct.

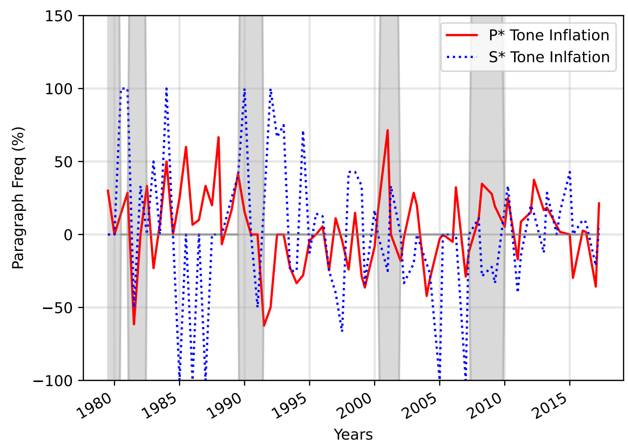
In particular, the risk measures from BERT, (3), have a much lower correlation at 0.18 which is not statistically significant. This is perhaps not surprising because the two measures likely confound risks associated with different topics. We explore this possibility further in section IV.D below, when we examine by topic the textual distance between the BERT measures more formally.



(A) Unemployment



(B) Real GDP Growth



(C) Inflation

Figure 9. Sentiment for Unemployment, GDP, and Inflation Topics as Classified by BERT

The figure displays the tone of GDP growth, unemployment, and inflation topics that are extracted using the BERT model. We show results for the FOMC (or P*) (red) and Greenbook (or S*) (blue) corpora. Shaded areas represent NBER recessions. The data are semiannual from 1979 through 2017.

Table 1—**Correlation of the Textual Measures**

FOMC	Greenbook	Correlation	P-value
P* Risk Uni	S* Risk Uni	0.61	0.00
P* Risk Bert	S* Risk Bert	0.18	0.11
P* Risk HHLT	S* Risk HHLT	0.28	0.01
P* Tone LM	S* Tone LM	0.48	0.00
P* Tone Bert	S* Tone Bert	0.34	0.00
P* Downside Risk	S* Downside Risk	0.39	0.00
P* Inflat Risk	S* Inflat Risk	0.17	0.13
P* Upside Risk	S* Upside Risk	0.36	0.00
P* Upside Risk Inflat	S* Equity Risk Premium	0.11	0.32
P* Downside Risk Growth	S* Risk Premium Corpor	0.18	0.13
P* Tone Unempl.	S* Tone Unempl.	0.20	0.08
P* Tone Inflation	S* Tone Inflation	0.06	0.62
P* Tone GDP	S* Tone GDP	0.14	0.23

Notes: The table presents correlation coefficients with corresponding p-values testing the significance of the correlation coefficient for each textual measure across the FOMC transcripts and the Greenbook. The underlying text-based measures are semiannual from 1979 to 2017.

IV. Text-Augmented Romer and Romer Regressions

A. *Empirical Setup*

Romer and Romer (2008) tested the information content of the FOMC's forecasts relative to those of the staff by estimating MZ regressions of the form:

$$X_t = a + bS_t + cP_t + e_t, \tag{6}$$

where X is the realized value of inflation, unemployment, or GDP growth, and S and P are the staff and policymaker (FOMC) quantitative forecasts for that variable. Such MZ regressions are a common way of evaluating the quality of forecasts under quadratic loss, as we discuss further below.

We follow Romer and Romer (2008) and estimate (6) pooled over the three forecast horizons. Since the forecasts are multi-step-ahead, we should not expect serial independence of e_t, even for well-calibrated point forecasts. When estimating (6) by OLS, we, therefore, compute Newey and West (1987) standard errors with three lags, the maximum lag at which one should expect serial correlation given the forecast horizon of our data introduced in Section II.

Our point of departure is to extend (6) so that we can also test the value of the text-based forecasts from S and P. This is achieved simply by augmenting (6) with the text-based measures considered above:

$$X_t = a + bS_t + cP_t + b^* S_t^* + c^* P_t^* + e_t, \tag{7}$$

where S_t^* and P_t^* are the text-based forecasts from S and P. These forecasts can be vectors, reflecting uncertainty about which text-based measures to consider. In effect, we can let the data determine which text-based measures (as available in real time) help explain X_t.

We start by deliberately pre-selecting the risk uni-gram as the text-based measure to consider. This is based on the view that discussions around "risk" are a principal way in which both the FOMC and Greenbook narratives communicate information about the expected quality and reliability of their point forecasts. Below we extend our analysis to consider the larger set of text-based measures discussed in Section III above, finding our main results, certainly for unemployment and GDP growth, to be robust.

Romer and Romer (2008) focused on estimating (6) by OLS. This regression is then familiar as the forecast combination regression of Granger and Ramanathan (1984), with the weights b and c indicating the "optimal" weight to attach to S and P under a quadratic loss function. Tests of $b = 0$ or $c = 0$ then amount to forecast "encompassing" tests, such that either S or P explains the predictive ability of its rival; for example, see Clements and Hendry (1998).

As Ellison and Sargent (2012) argue in their defense of the FOMC, FOMC and staff (point) forecasts may well be "answers to different questions." The FOMC may not be aiming to produce minimum mean squared error point forecasts, as implicitly assumed when we estimate (6) by OLS. Given model uncertainty, Ellison and Sargent propose a model whereby the FOMC optimally produces worst-case (biased in a statistical sense) forecasts. A related literature, while not testing the FOMC point forecasts *per se*, has found evidence that the staff (Greenbook) forecasts are better interpreted as intending to minimize an asymmetric rather than a quadratic loss function, where, for example, the costs of high inflation are greater than those of low inflation; see Capistrán (2008).

Under departures from quadratic loss, there is no reason that $b = 1$ or $c = 1$, even under forecast optimality of S or P. This is understood by noting that for a class of loss functions (homogeneous in the forecast error), the optimal forecast can be expressed as the conditional quantile (for example, see Patton and Timmermann (2007b)). To acknowledge this, and more generally model the conditional distribution of outcomes as a function of S, P, S^*, and P^*, we estimate (6) and, in turn, (7), as quantile regressions and thereby minimize the "generalized" forecast errors of Patton and Timmermann (2007a):

$$\min_{a,b,c} \sum_{t=1}^{T} \rho_\tau (X_t - a_\tau - b_\tau S_t - c_\tau P_t) \tag{8}$$

where

$$\rho_\tau(u) = u(\tau - I(u < 0))$$

is the check or asymmetric linear loss function of order τ, where $I(.)$ denotes an indicator function that places different weights on the errors depending on whether the error is above or below the τ-th quantile ($\tau \in (0, 1)$). An attraction of estimating quantile regressions is that we can test whether the forecasts, S, P, S^*, and P^* are useful not simply in explaining the conditional mean of X but also its distribution. S can be said to *encompass* P (see Giacomini and Komunjer (2005)) at quantile τ when $b_\tau = 1$ and $c_\tau = 0$.

B. The Role of Risk-Related Narratives

Unemployment. The results from estimating the MZ regressions, (6) and (7), by OLS and as quantile regressions for unemployment, GDP growth, and inflation are shown in Tables 2 – 4. Starting with unemployment in Table 2, and looking first at the OLS regression without P* and S*, we see that both the FOMC and the staff forecasts carry some weight. This result is, in fact, already less negative for the FOMC forecasts than when the same regression – as in Romer and Romer (2008) – is estimated over the shorter sample ending in 2001.[15] But when we add the text-based forecasts, P* and S*, we see two interesting features emerge. First, the information content of the FOMC point forecast rises. The FOMC point forecast encompasses the staff forecast (with a p-value of 0.00).[16] Second, the text-based measures themselves carry important informational content over and above that contained in P and S alone. Both risk-based measures, P* and S*, are informative about unemployment outcomes.

Turning to the quantile regressions, Table 2 shows that the FOMC point forecasts, P, are even

[15]We replicate this result from Romer and Romer (2008) in Table A15 in the online appendix, showing that for the sample ending in 2001 the staff forecasts for unemployment encompass those from the FOMC.

[16]This finding corroborates footnote 1 in Romer and Romer (2008), where Romer and Romer tentatively note, given their shorter sample that ends in 2001, that the informational content of FOMC forecasts appears higher when they reestimate (6) on a sample starting in 1990. On an extended sample through 2012, Binder and Wetzel (2018) also find evidence that FOMC point forecasts have improved since 2001. As in Table 2, Aruoba and Drechsel (2022) find that Greenbook text (S*) helps explain Greenbook unemployment rate forecast errors.

more informative in the upper tail, when unemployment is high. As we expand on below, this is consistent with the FOMC adopting a risk-management perspective when forming its forecasts. The estimated coefficient on P at the 80 percent quantile rises to 0.94 and that on S drops to 0.00. But, importantly, there is again statistically significant informational content in the text-based measures. Both FOMC and staff risk measures, P* and S*, are statistically significant. The offsetting signs on P* and S* can be understood by the relatively high correlation between these two measures. As shown in Table 1, P* and S* (when measured by the risk uni-grams) have a correlation coefficient of 0.61. This turns out to be a higher correlation coefficient than seen for any of the other text-based measures we consider, where there are clear differences between the FOMC and staff narrative measures for a given text-based algorithm.

Table 2—**Role of Staff and FOMC Forecasts and Narrative in Predicting Unemployment**

	(1) OLS	(2) OLS	(3) q20	(4) q20	(5) q80	(6) q80
P	0.61	0.84***	0.92**	0.70	0.94*	0.72
	(1.54)	(2.51)	(2.28)	(1.64)	(1.84)	(1.40)
S	0.30	0.13	-0.06	0.10	0.00	0.37
	(0.78)	(0.41)	(-0.16)	(0.25)	(0.01)	(0.68)
P*		1.80***		0.52		2.05**
		(2.94)		(1.59)		(1.97)
S*		-2.02***		-0.70**		-2.32**
		(-3.42)		(-2.18)		(-2.37)
Constant	0.55	-0.57	0.24	0.40	0.65	-0.80
	(1.62)	(-1.41)	(0.62)	(0.85)	(0.89)	(-1.02)
R^2	0.68	0.76	0.56	0.57	0.47	0.53
Obs	129	129	129	129	129	129
P* = 0		0.00		0.11		0.05
P* and P = 0	0.13	0.00	0.02	0.09	0.07	0.08
S* = 0		0.00		0.03		0.02
S* and S = 0	0.44	0.00	0.87	0.09	0.99	0.04

Notes: The dependent variable is the realized value of the variable being forecast. q20 and q80 refer to quantile regressions for the 20 percent and 80 percent quantiles. P* and S* measures are risk uni-grams. T-statistics are in parentheses; Newey and West (1987) standard errors used for OLS. * p < 0.10, ** p < 0.05, *** p < 0.01. Pseudo R^2 for the quantile regressions reported. P-values for the joint encompassing test that the coefficients on P^*, P, S^*, and/or S equal zero reported. The data are semiannual from 1979 to 2017.

GDP growth. Next we turn to GDP growth. Table 3 echoes the original result in Romer and Romer (2008), as replicated in online Table A16: FOMC forecasts have some useful information with a weight of 0.54 versus a weight of 0.22 for the staff. But neither forecast encompasses the other, consistent with Romer and Romer's argument that the FOMC and staff forecasts are pretty similar. Hence, the coefficients in (6) are not estimated with much precision. But, similarly to unemployment, adding in the text-based forecasts, P* and S*, changes what we characterize as the two main takeaways of Table 3. First, even in the extended OLS regression, column (2), we see that the FOMC point forecast encompasses the staff point forecast. Second, when we consider the text-augmented regression, P* is statistically significant in the OLS regression: the FOMC narrative around risk helps explain (mean) GDP outcomes over and above the point forecasts.

Turning to the quantile regressions, we see from Table 3 that the FOMC narrative adds information, especially in the upper tail. At the 80 percent quantile, we cannot reject the null hypothesis that there is no informational content to either the staff's point forecast or their narrative at the 5 percent significance level. The FOMC forecast and narrative encompass that of the staff. As indicated below, this result is weaker, but still evident when we consider a wider set of text-based measures.

Inflation. Table 4 turns to inflation. OLS estimation of (6) confirms that the "striking" result of Romer and Romer (2008) holds even on an extended sample through 2017. The staff point forecast for inflation encompasses the FOMC point forecast. The coefficient on the staff forecast is lower than in the original Romer and Romer sample: as shown in Table A17, it drops from 1.06, for the sample ending in 2001, to 0.65 in Table 4, suggesting that the FOMC forecast has become relatively more informative since 2001. But even so, the FOMC forecast remains statistically insignificant and is encompassed by the staff forecast. It is also noteworthy in Table 4 how the point forecasts for inflation explain considerably more of the variation in outcomes than do the unemployment or GDP forecasts: the R^2s are higher than in Tables 2 and 3. Put simply, the

TABLE 3—**Role of Staff and FOMC Forecasts and Narrative in Predicting GDP Growth**

	(1) OLS	(2) OLS	(3) q20	(4) q20	(5) q80	(6) q80
P	0.54	0.63*	0.63	1.00*	0.13	0.61
	(1.53)	(1.95)	(1.03)	(1.93)	(0.18)	(1.07)
S	0.22	0.17	0.41	-0.10	0.50	0.12
	(0.76)	(0.59)	(0.75)	(-0.20)	(0.84)	(0.25)
P*		-1.50***		-1.17**		-2.15***
		(-3.82)		(-2.22)		(-4.12)
S*		0.79		1.65**		0.49
		(1.46)		(1.98)		(0.72)
Constant	0.43	1.09***	-1.28**	-0.67	1.77***	2.64***
	(1.25)	(2.71)	(-2.21)	(-1.21)	(3.55)	(6.26)
R^2	0.36	0.43	0.25	0.30	0.20	0.30
Obs	129	129	129	129	129	129
P* = 0		0.00		0.03		0.00
P* and P = 0	0.13	0.00	0.29	0.03	0.88	0.00
S* = 0		0.15		0.05		0.47
S* and S = 0	0.45	0.30	0.46	0.14	0.40	0.72

Notes: The dependent variable is the realized value of the variable being forecast. q20 and q80 refer to quantile regressions for the 20 percent and 80 percent quantiles. P* and S* measures are risk uni-grams. T-statistics in parentheses; Newey and West (1987) standard errors used for OLS. * p < 0.10, ** p < 0.05, *** p < 0.01. Pseudo R^2 for the quantile regressions reported. P-values for the joint encompassing test that the coefficients on P^*, P, S^*, and/or S equal zero reported. The data are semiannual from 1979 to 2017.

inflation point forecasts appear better, perhaps because inflation is "easier" to forecast given its greater persistence. However, when we add in the text-based measures, we see that the FOMC, via its narrative, does in fact add value. But the value-added of the FOMC's risks narrative is weaker for inflation than for GDP and unemployment: P* is statistically significant only at the 10 percent level for the OLS regression in Table 4. The strength of the signal for inflation outcomes from both the FOMC and the staff narrative is higher, as summarized below, when we consider the wider set of text-based algorithms. For inflation, our results suggest that it is important to consult a wider set of narratives, perhaps because the narrative around inflation is more nuanced than what is captured *simply* by summing up the number of times "risk" is mentioned. This conclusion is also borne out when we look at the quantile regression results for inflation, to which we now turn.

Table 4 shows that the informational content of the FOMC point forecast is higher in the upper tail. The coefficient on P is higher, and more significant, at the 80 percent quantile than in the OLS regressions. This is again consistent with the model and findings in Ellison and Sargent (2012) that assume the FOMC adopts a risk-management perspective and does not seek to produce minimum mean squared error forecasts of inflation. The FOMC is more worried about high inflation outcomes. This also fits with the econometric results in Capistrán (2008), and with Binder and Wetzel (2018), who find that the relative forecasting performance of the FOMC is higher when economic conditions are worse. In contrast to our findings for unemployment and GDP growth, neither the FOMC nor the staff narrative about risk helps explain inflation outturns in the upper or lower tails. We now turn to an examination of whether other aspects of both narratives are informative by summarizing the estimation results of (7) when we consider our wider set of text-based measures.

C. Extended Regressions Using the Larger Set of Text-Based Factors

Tables 2 to 4 deliberately restrict attention to just one text-based measure: "risk," as measured via the uni-gram, (1). Now we add to (7) the full set of text-based factors introduced in Section III. Thereby, we test if there is additional value-added to be extracted from the FOMC and staff corpora if we consider P* and S* as vectors. But with more than 10 elements in both P* and S*, and given what remains relatively small sample sizes, we chose not to use OLS to estimate (7) with all 20 plus regressors. Instead, we estimate subset linear regressions, focusing on individual measures paired across S* and P*. The full estimation results across all pairs are presented in the online appendix (see Tables A3 – A14). When estimating the full regression with all the measures we use double Lasso (see Tibshirani (1996) and Belloni, Chernozhukov, and Hansen (2014)) to select the important variables.[17]

[17]Lasso estimates of the coefficients are based on minimization of the objective function: $\underset{\beta}{min}(Y - X\beta)'(Y - X\beta) + \lambda \sum_{j=1}^{p} |\beta_j|$, where p is the number of predictors. We denote by λ the regularization parameter that is selected via 10-fold cross-validation. The latter component of the objective function is the L1 penalty term.

TABLE 4—**Role of Staff and FOMC Forecasts and Narrative in Predicting Inflation**

	(1) OLS	(2) OLS	(3) q20	(4) q20	(5) q80	(6) q80
P	0.25	0.31	0.35	0.57	0.72	0.77
	(0.90)	(0.98)	(0.79)	(1.09)	(1.61)	(1.42)
S	0.65**	0.64**	0.46	0.28	0.26	0.26
	(2.21)	(1.99)	(1.26)	(0.61)	(0.63)	(0.57)
P*		0.46*		0.04		0.97
		(1.74)		(0.14)		(1.64)
S*		-0.01		0.37		-0.68
		(-0.03)		(0.84)		(-1.24)
Constant	0.25**	-0.16	0.07	-0.14	0.39***	-0.19
	(2.35)	(-0.71)	(0.29)	(-0.30)	(2.67)	(-0.48)
R^2	0.87	0.87	0.43	0.44	0.69	0.70
Obs	129	129	129	129	129	129
P* = 0		0.08		0.89		0.10
P* and P = 0	0.38	0.22	0.43	0.52	0.11	0.19
S* = 0		0.97		0.38		0.22
S* and S = 0	0.03	0.05	0.26	0.61	0.57	0.18

Notes: The dependent variable is the realized value of the variable being forecast. q20 and q80 refer to quantile regressions for the 20 percent and 80 percent quantiles. P* and S* measures are risk uni-grams. T-statistics in parentheses; Newey and West (1987) standard errors used for OLS. * p < 0.10, ** p < 0.05, *** p < 0.01. Pseudo R^2 for the quantile regressions reported. P-values for the joint encompassing test that the coefficients on P^*, P, S^*, and/or S equal zero reported. The data are semiannual from 1979 to 2017.

Table 5 – for unemployment – and Table 6 – for GDP – confirm that our earlier findings, seen in Tables 2 and 3, continue to hold when we consider the wider set of text-based measures. But we do see that there is additional information in some of the alternative text-based measures. The R^2 of the unemployment equation rises strongly from that in Table 2. For unemployment, FOMC references to "downside risk growth" are found to be especially important both for the OLS regression and in the tails of the distribution. For GDP, more of the text-based measures are important. Notably, staff assessments of "inflation risk" and FOMC talk of "downside growth risks" are statistically significant. While it makes sense to see upside inflation risks as informative about high GDP outcomes, it is perhaps a surprise to see talk of downside risks explaining upside outcomes. But we note the negative sign of the estimated coefficient in Table 6. The staff tone LM measure, (4), is also significant.

TABLE 5—**Predicting Unemployment with an Expanded Set of Text-Based Measures from the FOMC and the Staff**

	(1) OLS	(2) OLS	(3) q20	(4) q20	(5) q80	(6) q80
P	0.61	0.72***	0.92**	1.57***	0.94*	0.64**
	(1.54)	(2.65)	(2.28)	(3.94)	(1.84)	(2.17)
S	0.30	0.13	-0.06	-0.64*	0.00	0.30
	(0.78)	(0.51)	(-0.16)	(-1.74)	(0.01)	(0.92)
P* Tone LM		0.02		-0.04		-0.07
		(0.27)		(-0.51)		(-0.74)
S* Tone LM		0.12*		0.05		0.24**
		(1.68)		(0.47)		(2.47)
S* Tone Bert		0.02**		0.02		0.02
		(2.22)		(0.98)		(1.14)
P* Tone UNEM		0.00		0.00		0.01*
		(1.14)		(1.55)		(1.68)
S* Risk Bert		0.48		0.38		0.82**
		(1.48)		(0.98)		(2.05)
S* Risk HHLT		-0.77***		-0.30*		-0.90***
		(-4.45)		(-1.66)		(-4.15)
P* Downsid Risk Growth		43.30***		25.10**		70.76***
		(3.38)		(2.35)		(3.37)
S* Risk Premium Corpor		24.31***		19.80		16.15
		(3.34)		(1.40)		(0.90)
Constant	0.55	0.60*	0.24	-0.02	0.65	0.44
	(1.62)	(1.87)	(0.62)	(-0.05)	(0.89)	(0.63)
R^2	0.68	0.83	0.56	0.63	0.47	0.66
Obs	129	129	129	129	129	129
P* = 0		0.00		0.04		0.00
P*= 0 and P = 0	0.13	0.00	0.02	0.00	0.07	0.00
S* = 0		0.00		0.11		0.00
S* = 0 and S = 0	0.44	0.00	0.87	0.11	0.99	0.00

Notes: The dependent variable is the realized value of the variable being forecast. q20 and q80 refer to quantile regressions for the 20 percent and 80 percent quantiles. P* and S* measures are selected from the full set of measures by double Lasso. T-statistics in parentheses; Newey and West (1987) standard errors used for OLS. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Pseudo R^2 for the quantile regressions reported. The data are semiannual from 1979 to 2017.

For inflation, as anticipated above, we see that it is even more important to look beyond the risk uni-grams when aiming to capture the forecasting information in the narratives from the FOMC and the staff. Table 7 shows that for inflation a greater number of the text-based measures are selected. As a result, for inflation in assessing the relative informational content of the FOMC and the staff, and of their point forecasts versus their narrative, it does matter

TABLE 6—**Predicting GDP Growth with an Expanded Set of Text-Based Measures from the FOMC and the Staff**

	(1) OLS	(2) OLS	(3) q20	(4) q20	(5) q80	(6) q80
P GDP	0.54	0.66***	0.63	0.66***	0.13	0.58***
	(1.53)	(5.85)	(1.06)	(3.08)	(0.18)	(4.66)
S GDP	0.22		0.41		0.50	
	(0.76)		(0.75)		(0.84)	
S* Tone LM		-0.33***		-0.59***		-0.23
		(-3.52)		(-3.24)		(-1.28)
S* Inflat Risk		-9.15		-0.56		-23.92**
		(-1.50)		(-0.05)		(-2.10)
P* Upsid Risk Inflat		-38.12*		-83.51**		-24.03
		(-1.93)		(-2.08)		(-1.38)
P* Downsid Risk Growth		-27.58***		-10.78		-39.80**
		(-2.76)		(-0.52)		(-2.42)
S* Risk Premium Corpor		-12.47		-6.98		-16.16
		(-1.13)		(-0.24)		(-1.11)
Constant	0.43	2.22***	-1.28**	2.27*	1.77***	3.02***
	(1.25)	(4.29)	(-2.21)	(1.89)	(3.55)	(4.02)
R^2	0.36	0.52	0.25	0.38	0.20	0.33
Obs	129	129	129	129	129	129
P* = 0		0.00		0.00		0.00
P*= 0 and P = 0	0.13	0.00	0.29	0.00	0.86	0.00
S* = 0		0.00		0.01		0.02
S* = 0 and S = 0	0.45	0.00	0.46	0.01	0.40	0.02

Notes: The dependent variable is the realized value of the variable being forecast. q20 and q80 refer to quantile regressions for the 20 percent and 80 percent quantiles. P* and S* measures are selected from the full set of measures by double Lasso. T-statistics in parentheses; Newey and West (1987) standard errors used for OLS. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Pseudo R^2 for the quantile regressions reported. The data are semiannual from 1979 to 2017.

whether we consult Table 4 or Table 7. Table 7 shows that both the FOMC and the staff narrative, as captured by the wider set of measures of P* and S*, is informative in the OLS regressions. Interestingly, the FOMC point forecast is statistically significant, while the staff forecast is not, once we control for P* and S*. It is also noteworthy that it is in fact the FOMC risk uni-gram that is the most informative single measure. But this is only revealed when we do condition on the wider set of text-based measures. For inflation, more of the text-based measures appear to help predict the realizations, albeit they have quite small estimated weights in the MZ regression.

While Table 7 confirms the results of Table 4 for the lower quantile – that both FOMC and staff point and narrative forecasts are uninformative for inflation – for the upper quantile we see the informational content of both the FOMC point forecast and narrative rise. P and P* encompass S and S*. P receives a weight of unity and is statistically significant, while S receives a small and statistically insignificant weight.

Overall, therefore, our results demonstrate that there is value-added in the FOMC narrative, in particular when forecasting the tails of the outcome distribution. The FOMC narrative is both distinct, as we explore further in the next section, from the narrative in the Greenbook and it contains information over and above that captured by the point forecasts.[18]

D. Textual Similarity Between the FOMC and the Greenbook

In this section, having estimated the BERT model on each corpus, we more formally contrast the informational content of the FOMC and Greenbook corpora. The similarity metric used requires the same number of paragraphs in each corpus. So we focus on paragraphs that belong to each BERT topic, selecting the top paragraphs based on their probability.[19]

Specifically, we construct textual "distance" measures between the BERT-based P* and S* risk, unemployment, GDP growth, and inflation topics. We select the top 20 paragraphs from each corpus with the highest probability of discussing each topic. Then, we calculate the cosine similarity between the top 20 paragraphs between the FOMC and Greenbook for each topic.[20] This method has been used in other studies; for example, see Hoberg and Phillips (2016).[21]

The cosine similarity measure captures the cosine of the angle between two n-dimensional vectors projected in a multi-dimensional space. The two vectors correspond to the FOMC and Greenbook corpora. The cosine similarity of the two corpora is bounded between 0 and 1. A

[18] In the online appendix (see Section A2), we show that these main results are generally robust to possible temporal instabilities and nonlinearities in the Romer and Romer regressions.

[19] This exercise is more challenging for the bag-of-words approach, as there is no probability assigned to each paragraph, making it hard to rank them. Hence our more formal analysis of textual similarity is confined to output from BERT.

[20] The selection of the number of paragraphs does not affect our results.

[21] Gentzkow, Kelly, and Taddy (2019) offer an excellent review of the literature.

TABLE 7—**Predicting Inflation with an Expanded Set of Text-Based Measures from the FOMC and the Staff**

	(1) OLS	(2) OLS	(3) q20	(4) q20	(5) q80	(6) q80
P	0.25	0.53*	0.35	0.66	0.72	1.03*
	(0.87)	(1.77)	(0.90)	(1.55)	(1.61)	(1.89)
S	0.65**	0.49	0.46	0.27	0.26	0.11
	(2.21)	(1.65)	(1.26)	(0.71)	(0.57)	(0.20)
P* Tone LM		-0.13*		-0.08		-0.06
		(-1.74)		(-0.90)		(-0.66)
S* Tone LM		-0.07		-0.06		-0.17
		(-0.67)		(-0.51)		(-1.13)
S* Tone Inflation		0.21		0.00		0.09
		(1.16)		(0.88)		(0.34)
P* Risk Bert		7.34		0.12		-2.73
		(1.00)		(1.34)		(-0.19)
P* Risk Uni		0.91**		0.59		1.89**
		(2.33)		(1.14)		(2.03)
P* Risk HHLT		-0.47**		-0.45		-0.56*
		(-2.00)		(-1.14)		(-1.85)
S* Risk HHLT		-0.20		0.13		-0.35
		(-1.01)		(0.57)		(-0.98)
S* Upsid Risk		-5.53***		-2.76		-9.17*
		(-2.76)		(-0.89)		(-1.71)
S* Equiti Risk Premium		13.55**		4.67		13.01
		(2.20)		(0.53)		(0.98)
Constant	0.25**	0.79**	0.07	0.22	0.39***	0.79
	(2.35)	(2.24)	(0.28)	(0.41)	(2.67)	(0.37)
R^2	0.87	0.89	0.43	0.48	0.69	0.73
Obs	129	129	129	129	129	129
P* = 0		0.00		0.19		0.02
P*= 0 and P = 0	0.38	0.00	0.37	0.24	0.11	0.04
S* = 0		0.02		0.70		0.24
S* = 0 and S = 0	0.03	0.00	0.21	0.76	0.57	0.08

Notes: The dependent variable is the realized value of the variable being forecast. q20 and q80 refer to quantile regressions for the 20 percent and 80 percent quantiles. P* and S* measures are selected from the full set of measures by double Lasso. T-statistics in parentheses; Newey and West (1987) standard errors used for OLS. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Pseudo R^2 for the quantile regressions reported. The data are semiannual from 1979 to 2017.

cosine similarity of 1 implies that the two documents have the same orientation, and if the measure is close to 0, the two corpora have a smaller similarity. Thus, the cosine similarity ($cos(P^*_{\text{TF-IDF}}, S^*_{\text{TF-IDF}})$) is defined as:

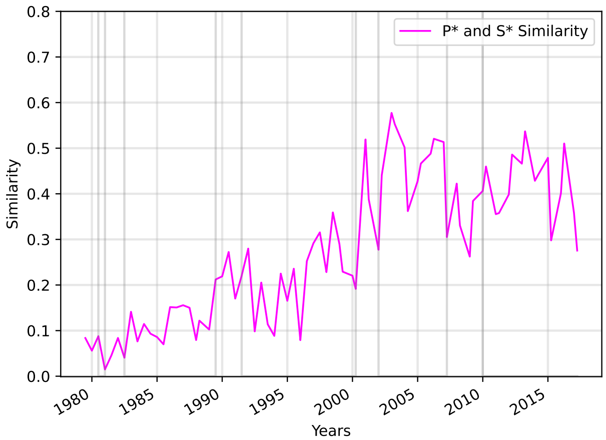
$$cos\left(P^{*}_{\text{TF-IDF}}, S^{*}_{\text{TF-IDF}}\right) = \frac{\sum_{i=1}^{N} P^{*}_{i,\text{TF-IDF}} S^{*}_{i,\text{TF-IDF}}}{\sqrt{\sum_{i=1}^{N} P^{*2}_{i,\text{TF-IDF}}}\sqrt{\sum_{i=1}^{N} S^{*2}_{i,\text{TF-IDF}}}}, \tag{9}$$

where $P^{*}_{\text{TF-IDF}}$ and $S^{*}_{\text{TF-IDF}}$ denote the TF-IDF vectors. Term frequency (TF) denotes the number of times a word appears in a paragraph over the total number of words in the paragraph. The inverse document frequency (IDF) is defined as $log(N/n)$, where N represents the total number of paragraphs in each corpus and n denotes the number of paragraphs with the specific term. Thus, the TF-IDF metric for each word is computed as the product of the two terms.

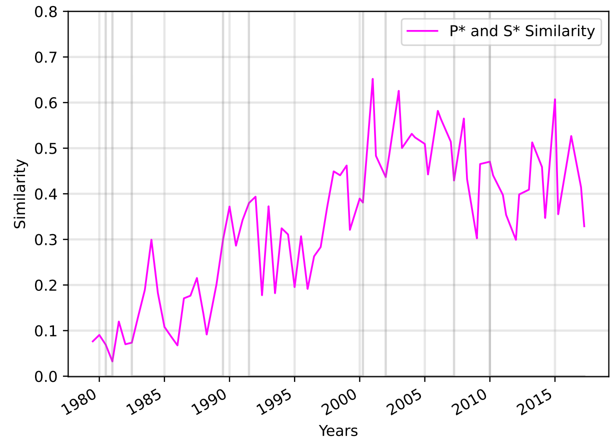
Figure 10 shows the cosine similarity between the FOMC transcripts and the Greenbook for the unemployment, real GDP growth, inflation, and risk topics over time. We find that similarity increases during financial crises, with the exception of the global financial crisis (GFC) of 2008-2009. This is perhaps not surprising, because there were strong differences in opinion and a diversity of views during the GFC about how long the recession would last and what the response of the central bank should be. This could explain higher disagreement during this period. For example, it was unclear what the impact of quantitative easing (QE) would be. This finding is consistent across the different topics. Another important finding is that the similarity between the two corpora increases over time. This positive trend implies some convergence in opinion between the staff and the FOMC. As we highlighted earlier, the only exception is during the GFC, when we observe a decline in similarity. It is worth noting that the trend is less steep for the risk topic in comparison to other topics, and the similarity level is low, on average, which illustrates the persistent disagreement between the FOMC transcripts and Greenbook regarding the different types of risks in the economy. Specifically, the maximum level of similarity for the risk topic is around 40 percent, while for the other topics it ranges from 55-65 percent, with the greatest similarity for real GDP growth.

Overall, we find that the similarity between the two corpora for different topics is low, on average, even though it tends to be increasing over time. This helps explain why in the text-augmented Romer and Romer regression we found it important to consult both the FOMC and

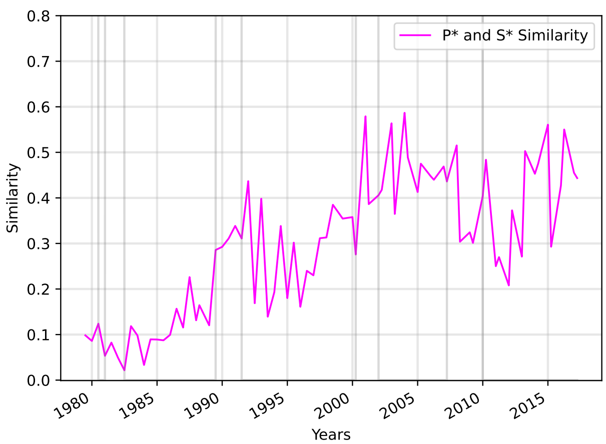
Greenbook documents, as they provide distinct information. Specifically, Figure 10 illustrates that their information content is different, and policymakers should take into account the content of both corpora as also emphasized in our extended Romer and Romer regressions.



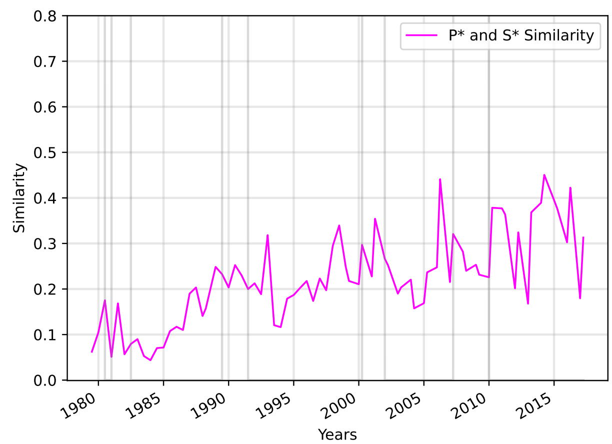
(A) Unemployment



(B) Real GDP Growth



(C) Inflation



(D) Risk

Figure 10. Cosine Similarity between S* and P* for the Unemployment, GDP, and Inflation Topics as Classified by BERT

The figure displays the cosine similarity between S* and P* for the GDP growth, unemployment, and inflation topics estimated using BERT. The data are semiannual from 1979 through 2017.

V. Conclusions

This paper extends Romer and Romer (2008) to test the value, relative to Board of Governors staff, of not just US monetary policymakers' quantitative forecasts but their qualitative assessments of the state of the economy and its risks and uncertainties. Using a range of methods from natural language processing, we find that the FOMC does add value – via its words, its tales – when forecasting. Information in the FOMC narrative about macroeconomic risks and uncertainties is especially helpful in explaining tail outcomes. Our results are consistent with evidence, as formalized by Ellison and Sargent (2012), that FOMC members tend to follow a risk-management perspective when producing their point forecasts. We find that in producing both its quantitative (point) forecasts and its accompanying narrative, the FOMC is attuned to macroeconomic risks and uncertainties.

Our results also have direct implications for the identification of monetary policy shocks via the narrative approach of Romer and Romer (2004). Aruoba and Drechsel (2022) show that information in the Greenbook (and other documents prepared by the staff, including the Redbook/Beigebook) helps explain movements in the federal funds rate beyond the quantitative forecasts prepared by the staff. This means that the exogenous component of monetary policy is much smaller than if information in only the quantitative forecasts is entertained. The results in this paper imply that, when identifying monetary policy shocks: (i) it is improper not to distinguish between the information sets of the staff and policymakers, since our text-based analysis reveals clear differences between their narratives,[22] and (ii) it is important to model the entire distribution of outcomes not just the conditional mean.

[22] Table A18 in the online appendix corroborates this statement: FOMC, rather than staff, narratives about "risk" are seen to improve the fit of a Romer and Romer (2004) regression of the change in the intended federal funds rate on P's and S's point forecasts of GDP growth, inflation, and the unemployment rate. Future research should extend this illustrative result to consider both higher-frequency data and a greater range of textual and point forecasts following Aruoba and Drechsel (2022), importantly differentiating between the tales of the FOMC and the staff.

References

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone (2019). "Vulnerable growth." *American Economic Review*, 109(4), pp. 1263–89. doi:10.1257/aer.20161923.

Arai, Natsuki (2016). "Evaluating the efficiency of the FOMC's new economic projections." *Journal of Money, Credit and Banking*, 48(5), pp. 1019–1049. doi:10.1111/jmcb.12325.

Aruoba, Boragan and Thomas Drechsel (2022). "Identifying monetary policy shocks: A natural language approach." CEPR Discussion Paper 17133. URL https://ideas.repec.org/p/cpr/ceprdp/17133.html.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2016). "Measuring economic policy uncertainty." *The Quarterly Journal of Economics*, 131(4), pp. 1593–1636. doi:10.1093/qje/qjw024.

Balke, Nathan S., Michael Fulmer, and Ren Zhang (2017). "Incorporating the Beige Book into a quantitative index of economic activity." *Journal of Forecasting*, 36(5), pp. 497–514. doi:10.1002/for.2450.

Banternghansa, Chanont and Michael W. McCracken (2009). "Forecast disagreement among FOMC members." Federal Reserve Bank of St. Louis Working Paper 2009-059A. doi:10.20955/wp.2009.059.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2014). "Inference on treatment effects after selection among high-dimensional controls." *The Review of Economic Studies*, 81(2), pp. 608–650. doi:10.1093/restud/rdt044.

Binder, Carola Conces and Samantha Wetzel (2018). "The FOMC versus the staff, revisited: When do policymakers add value?" *Economics Letters*, 171, pp. 72–75. doi:10.1016/j.econlet.2018.07.006.

Capistrán, Carlos (2008). "Bias in Federal Reserve inflation forecasts: Is the Federal Reserve irrational or just cautious?" *Journal of Monetary Economics*, 55(8), pp. 1415–1427. doi:10.1016/j.jmoneco.2008.09.011.

Chava, Sudheer, Wendi Du, and Baridhi Malakar (2021). "Do managers walk the talk on environmental and social issues?" Technical Report 3900814, Georgia Tech Scheller College of Business Research Paper. doi:10.2139/ssrn.3900814.

Clements, Michael and David Hendry (1998). *Forecasting Economic Time Series*. Cambridge University Press. doi:10.1017/CBO9780511599286.

Clements, Michael P. and J. James Reade (2020). "Forecasting and forecast narratives: The Bank of England Inflation Reports." *International Journal of Forecasting*, 36(4), pp. 1488–1500. doi:10.1016/j.ijforecast.2019.08.013.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2018). "BERT: Pre-training of deep bidirectional transformers for language understanding." Technical report, arXiv:1810.04805. doi:10.48550/arXiv.1810.04805.

Ellison, Martin and Thomas J. Sargent (2012). "A defense of the FOMC." *International Economic Review*, 53(4), pp. 1047–1065. doi:10.1111/j.1468-2354.2012.00711.x.

Faust, Jon and Jonathan H. Wright (2009). "Comparing Greenbook and reduced form forecasts using a large realtime dataset." *Journal of Business and Economic Statistics*, 27(4), pp. 468–479. doi:10.1198/jbes.2009.07214.

Filippou, Ilias, My T. Nguyen, and Ganesh Viswanath-Natraj (2023). "Fundamental vs. technical analysis: News-based factors and and cryptocurrency risk premia." Technical report, Available at SSRN 4402365. doi:10.2139/ssrn.4402365.

Gardner, Ben, Chiara Scotti, and Clara Vega (2022). "Words speak as loudly as actions: Central bank communication and the response of equity prices to macroeconomic announcements." *Journal of Econometrics*, 231(2), pp. 387–409. doi:10.1016/j.jeconom.2021.07.014.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy (2019). "Text as data." *Journal of Economic Literature*, 57(3), pp. 535–574. doi:10.1257/jel.20181020.

Giacomini, Raffaella and Ivana Komunjer (2005). "Evaluation and combination of conditional quantile forecasts." *Journal of Business and Economic Statistics*, 23(4), pp. 416–431. doi:10.1198/073500105000000018.

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera (2023). "The voice of monetary policy." *American Economic Review*, 113(2), pp. 548–84. doi:10.1257/aer.20220129.

Granger, Clive W. J. and Ramu Ramanathan (1984). "Improved methods of combining forecasts." *Journal of Forecasting*, 3(2), pp. 197–204. doi:10.1002/for.3980030207.

Hansen, Stephen and Michael McMahon (2016). "Shocking language: Understanding the macroeconomic effects of central bank communication." *Journal of International Economics*, 99, pp. S114–S133. doi:10.1016/j.jinteco.2015.12.008.

Hansen, Stephen, Michael McMahon, and Andrea Prat (2018). "Transparency and deliberation within the FOMC: A computational linguistics approach." *Quarterly Journal of Economics*, 133(2), pp. 801–870. doi:10.1093/qje/qjx045.

Hansen, Stephen, Michael McMahon, and Matthew Tong (2019). "The long-run information effect of central bank communication." *Journal of Monetary Economics*, 108, pp. 185–202. doi:10.1016/j.jmoneco.2019.09.002.

Harris, Zellig S. (1954). "Distributional structure." *Word*, 10(2-3), pp. 146–162. doi:10.1080/00437956.1954.11659520.

Hassan, Tarek A, Stephan Hollander, Laurence Van Lent, and Ahmed Tahoun (2019). "Firm-level political risk: Measurement and effects." *Quarterly Journal of Economics*, 134(4), pp. 2135–2202. doi:10.1093/qje/qjz021.

Hoberg, Gerard and Gordon Phillips (2016). "Text-based network industries and endogenous product differentiation." *Journal of Political Economy*, 124(5), pp. 1423–1465. doi:10.1086/688176.

Joulin, Armand, Edouard Grave, Piotr Bojanowski, Matthijs Douze, Hérve Jégou, and Tomas Mikolov (2016). "FastText.zip: Compressing text classification models." Technical report, arXiv:1612.03651. doi:10.48550/arXiv.1612.03651.

Kalamara, Eleni, Arthur Turrell, Chris Redl, George Kapetanios, and Sujit Kapadia (2022). "Making text count: Economic forecasting using newspaper text." *Journal of Applied Econometrics*, 37(5), pp. 896–919. doi:10.1002/jae.2907.

Lima, Luiz Renato, Lucas Lúcio Godeiro, and Mohammed Mohsin (2021). "Time-varying dictionary and the predictive power of FED minutes." *Computational Economics*, 57(1), pp. 149–181. doi:10.1007/s10614-020-10039-9.

Loughran, Tim and Bill McDonald (2011). "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks." *The Journal of Finance*, 66(1), pp. 35–65. doi:10.1111/j.1540-6261.2010.01625.x.

Mikolov, Tomas, Kai Chen, Gregory S. Corrado, and Jeffrey A. Dean (2017). "Computing numeric representations of words in a high-dimensional space." US Patent 9,740,680. URL https://patentimages.storage.googleapis.com/a5/7f/96/36fe34de5ae176/US9740680.pdf.

Mincer, Jacob and Victor Zarnowitz (1969). "The evaluation of economic forecasts." In Jacob Mincer, editor, *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and*

Performance, pp. 3–46. National Bureau of Economic Research, Inc. URL https://www.nber.org/system/files/chapters/c1214/c1214.pdf.

Newey, Whitney K. and Kenneth D. West (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica*, 55(3), pp. 703–708. doi:10.2307/1913610.

Patton, Andrew J. and Allan Timmermann (2007a). "Properties of optimal forecasts under asymmetric loss and nonlinearity." *Journal of Econometrics*, 140(2), pp. 884 – 918. doi:10.1016/j.jeconom.2006.07.018.

Patton, Andrew J. and Allan Timmermann (2007b). "Testing forecast optimality under unknown loss." *Journal of the American Statistical Association*, 102(480), pp. 1172–1184. doi:10.1198/016214506000001176.

Peek, Joe, Eric Rosengren, and Geoffrey M. B. Tootell (2016). "Does Fed policy reveal a ternary mandate?" Working Papers 16-11, Federal Reserve Bank of Boston. URL https://ideas.repec.org/p/fip/fedbwp/16-11.html.

Pennington, Jeffrey, Richard Socher, and Christopher D. Manning (2014). "GloVe: Global vectors for word representation." In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543. doi:10.3115/v1/D14-1162.

Perone, Christian S., Roberto Silveira, and Thomas S. Paula (2018). "Evaluation of sentence embeddings in downstream and linguistic probing tasks." Technical report, arXiv:1806.06259. doi:10.48550/arXiv.1806.06259.

Porter, Martin F. (1980). "An algorithm for suffix stripping." *Program*, 14(3), pp. 130–137. doi:10.1108/eb046814.

Romer, Christina D. and David H. Romer (2000). "Federal Reserve information and the behavior of interest rates." *American Economic Review*, 90(3), pp. 429–457. doi:10.1257/aer.90.3.429.

Romer, Christina D. and David H. Romer (2004). “A new measure of monetary shocks: Derivation and implications.” *American Economic Review*, 94(4), pp. 1055–1084. doi:10.1257/0002828042002651.

Romer, Christina D. and David H. Romer (2008). “The FOMC versus the staff: Where can monetary policymakers add value?” *American Economic Review*, 98(2), pp. 230–35. doi:10.1257/aer.98.2.230.

Romer, David (2010). “A new data set on monetary policy: The economic forecasts of individual members of the FOMC.” *Journal of Money, Credit and Banking*, 42(5), pp. 951–957. doi:10.1111/j.1538-4616.2010.00314.x.

Schmanski, Bennett, Chiara Scotti, Clara Vega, and Hedi Benamar (2023). “Fed communication, news, Twitter, and echo chambers.” Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System (U.S.). doi:10.17016/FEDS.2023.036.

Segal, Gill, Ivan Shaliastovich, and Amir Yaron (2015). “Good and bad uncertainty: Macroeconomic and financial market implications.” *Journal of Financial Economics*, 117(2), pp. 369–397. doi:10.1016/j.jfineco.2015.05.004.

Shapiro, Adam Hale and Daniel J. Wilson (2022). “Taking the Fed at its word: A new approach to estimating central bank objectives using text analysis.” *The Review of Economic Studies*, 89(5), pp. 2768–2805. doi:10.1093/restud/rdab094.

Shapley, Lloyd S. (1953). “A value for n-person games.” *Contributions to the Theory of Games*, 2(28), pp. 307–317. URL https://www.degruyter.com/document/doi/10.1515/9781400881970-018/html.

Sharpe, Steven A., Nitish R. Sinha, and Christopher A. Hollrah (2023). “The power of narrative sentiment in economic forecasts.” *International Journal of Forecasting*, 39(3), pp. 1097–1121. doi:10.1016/j.ijforecast.2022.04.008.

Shiller, Robert J. (2017). "Narrative economics." *American Economic Review*, 107(4), pp. 967–1004. doi:10.1257/aer.107.4.967. URL https://www.aeaweb.org/articles?id=10.1257/aer.107.4.967.

Sims, Christopher (2002). "The role of models and probabilities in the monetary policy process." *Brookings Papers on Economic Activity*, 33(2), pp. 1–62. doi:10.1353/ECA.2003.0009.

Stekler, Herman and Hilary Symington (2016). "Evaluating qualitative forecasts: The FOMC minutes, 2006-2010." *International Journal of Forecasting*, 32(2), pp. 559–570. doi:10.1016/j.ijforecast.2015.02.003.

Strumbelj, Erik and Igor Kononenko (2010). "An efficient explanation of individual classifications using game theory." *Journal of Machine Learning Research*, 11(1), pp. 1–18. URL https://www.jmlr.org/papers/volume11/strumbelj10a/strumbelj10a.pdf.

Tetlock, Paul C. (2007). "Giving content to investor sentiment: The role of media in the stock market." *The Journal of Finance*, 62(3), pp. 1139–1168. doi:10.1111/j.1540-6261.2007.01232.x.

Tibshirani, Robert (1996). "Regression shrinkage and selection via the Lasso." *Journal of the Royal Statistical Society Series B (Methodological)*, 58(1), pp. 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x.

Yang, Yi, Mark Christopher Siy Uy, and Allen Huang (2020). "FinBERT: A pretrained language model for financial communications." doi:10.48550/arXiv.2006.08097.

Online Appendix for

“The FOMC versus the Staff: Do Policymakers Add Value in Their Tales?”

by

Ilias Filippou James Mitchell My T. Nguyen

A1. Additional Tables and Figures as Referenced in the Main Paper[23]

[23]The tables and figures in this appendix make use of notation as described in the main paper.

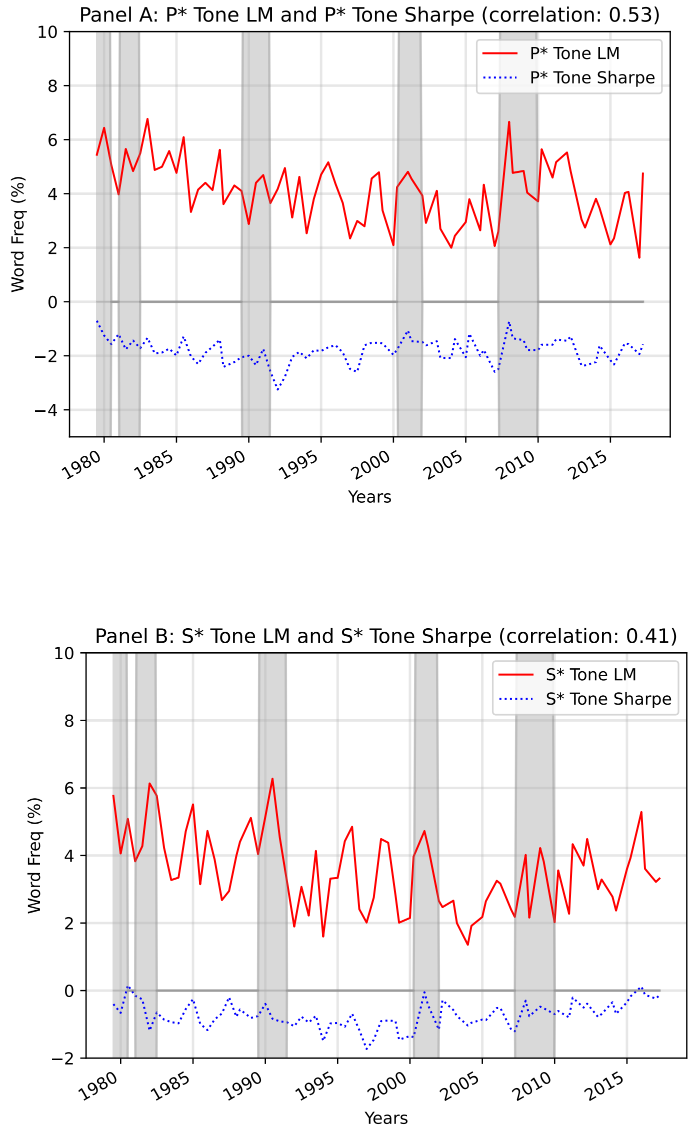
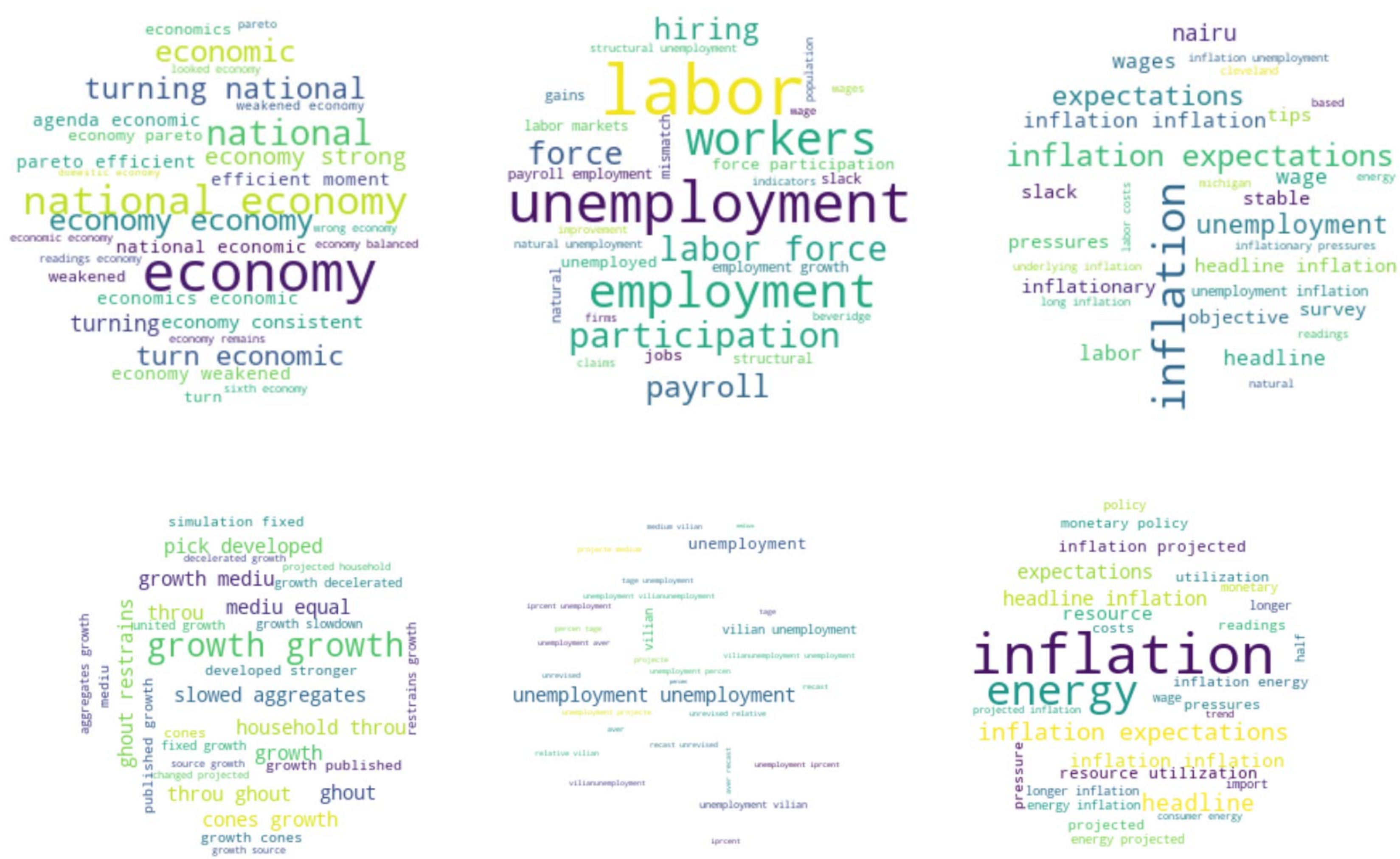


Figure A1. Tone LM and Tone Sharpe

The figure displays tone LM and tone Sharpe. We show results for FOMC (or P*) (top) and Greenbook (or S*) (bottom) corpora. The data are semiannual from 1979 to 2017.

Figure A2. Text-Based Index (Risk Attention and Tone)



The figure displays BERT topic word clouds for three topics. We show results for FOMC (or P*) (top) and Greenbook (or S*) (bottom) corpora. The data are from 1979:Q4 to 2017:Q2.

TABLE A1—**Bi-grams and Tri-grams**

The table presents the frequency of the top 10 bi-grams with the words "risk" in the FOMC transcripts and in the Greenbook in percentage points. The data are semiannual from 1979 through 2017.

Panel A: FOMC Transcripts Bi-grams

Terms	Term Frequency
P* downsid_risk	0.056
P* risk_inflat	0.027
P* balanc_risk	0.025
P* inflat_risk	0.021
P* upsid_risk	0.019
P* risk_premium	0.014
P* risk_manag	0.010
P* risk_growth	0.011
P* risk_balanc	0.010
P* assess_risk	0.010

Panel B: FOMC Transcripts Tri-grams

Terms	Term Frequency
P* upsid_risk_inflat	0.004
P* downsid_risk_growth	0.004
P* risk_financi_stabil	0.003

Panel C: Greenbook Bi-grams

Terms	Term Frequency
S* risk_premium	0.020
S* downsid_risk	0.010
S* upsid_risk	0.008
S* risk_spread	0.007
S* view_risk	0.003
S* illustr_risk	0.004
S* inflat_risk	0.004
S* equiti_risk	0.003
S* macroeconom_risk	0.002

Panel D: Greenbook Tri-grams

Terms	Term Frequency
S* equiti_risk_premium	0.003
S* risk_premium_corpor	0.003
S* inflat_risk_premium	0.002

Table A2—Correlation coefficients between the different text-based factors within the FOMC (Panel A) and Greenbook (Panel B) corpora; p-values for statistical significance in parentheses

Panel A: FOMC Transcripts

Variables	P* Risk Uni-gram	P* Risk BERT	P* Risk HHLT	P* Tone LM	P* Tone BERT	P* Downside Risk	P* Inflat Risk	P* Upsid Risk	P* Upside Risk Inflat	P* Downside Risk Growth	P* Tone Unemploy	P* Tone Inflation	P* Tone GDP
P* Risk Uni-gram	1.00												
P* Risk BERT	0.88	1.00											
	(0.00)												
P* Risk HHLT	0.64	0.56	1.00										
	(0.00)	(0.00)											
P* Tone LM	-0.12	-0.18	0.17	1.00									
	(0.28)	(0.11)	(0.14)										
P* Tone Bert	0.05	-0.02	0.20	0.70	1.00								
	(0.67)	(0.84)	(0.08)	(0.00)									
P* Downside Risk	0.75	0.68	0.59	0.15	0.24	1.00							
	(0.00)	(0.00)	(0.00)	(0.18)	(0.03)								
P* Inflat Risk	0.53	0.45	0.28	-0.14	-0.12	0.35	1.00						
	(0.00)	(0.00)	(0.01)	(0.22)	(0.31)	(0.00)							
P* Upside Risk	0.51	0.53	0.38	-0.27	-0.20	0.41	0.43	1.00					
	(0.00)	(0.00)	(0.00)	(0.02)	(0.08)	(0.00)	(0.00)						
P* Upside Risk Inflat	0.59	0.54	0.41	-0.06	-0.02	0.46	0.53	0.73	1.00				
	(0.00)	(0.00)	(0.00)	(0.62)	(0.85)	(0.00)	(0.00)	(0.00)					
P* Downside Risk Growth	0.68	0.52	0.48	0.12	0.17	0.60	0.55	0.44	0.69	1.00			
	(0.00)	(0.00)	(0.00)	(0.31)	(0.13)	(0.00)	(0.00)	(0.00)	(0.00)				
P* Tone Unemployment	0.19	0.13	0.08	0.24	0.29	0.29	0.09	0.02	0.07	0.23	1.00		
	(0.10)	(0.25)	(0.47)	(0.04)	(0.01)	(0.01)	(0.45)	(0.89)	(0.56)	(0.05)			
P* Tone Inflation	0.09	0.03	0.27	0.37	0.48	0.18	-0.01	-0.16	-0.01	0.12	0.09	1.00	
	(0.43)	(0.80)	(0.02)	(0.00)	(0.00)	(0.12)	(0.91)	(0.17)	(0.92)	(0.32)	(0.43)		
P* Tone GDP	-0.15	-0.20	0.09	0.39	0.55	0.07	-0.26	-0.21	-0.23	-0.07	0.24	0.39	1.00
	(0.20)	(0.08)	(0.44)	(0.00)	(0.00)	(0.56)	(0.02)	(0.07)	(0.05)	(0.53)	(0.03)	(0.00)	

Panel B: Greenbook

Variables	S* Risk Uni-gram	S* Risk BERT	S* Risk HHLT	S* Tone LM	S* Tone BERT	S* Downsid Risk	S* Inflat Risk	S* Upside Risk	S* Equity Risk Premium	P* Risk Premium Corpor	S* Tone Unemploy	S* Tone GDP	S* Tone Inflation
S* Risk Uni-gram	1.00												
S* Risk BERT	0.34	1.00											
	(0.00)												
S* Risk HHLT	0.89	0.30	1.00										
	(0.00)	(0.01)											
S* Tone LM	-0.12	-0.02	-0.02	1.00									
	(0.31)	(0.87)	(0.87)										
S* Tone BERT	0.13	0.03	0.15	0.61	1.00								
	(0.25)	(0.81)	(0.18)	(0.00)									
S* Downside Risk	0.65	0.12	0.59	-0.08	0.06	1.00							
	(0.00)	(0.28)	(0.00)	(0.48)	(0.59)								
S* Inflat Risk	0.48	0.18	0.48	-0.09	0.04	0.34	1.00						
	(0.00)	(0.12)	(0.00)	(0.44)	(0.71)	(0.00)							
S* Upside Risk	0.14	0.04	0.05	-0.21	-0.08	0.23	0.20	1.00					
	(0.23)	(0.75)	(0.69)	(0.06)	(0.51)	(0.05)	(0.09)						
S* Equity Risk Premium	0.48	0.47	0.36	-0.00	0.09	0.09	0.16	-0.05	1.00				
	(0.00)	(0.00)	(0.00)	(0.98)	(0.42)	(0.44)	(0.15)	(0.65)					
P* Risk Premium Corpor	0.37	0.22	0.23	-0.17	0.01	0.06	0.25	0.21	0.63	1.00			
	(0.00)	(0.06)	(0.04)	(0.14)	(0.91)	(0.63)	(0.03)	(0.07)	(0.00)				
S* Tone Unemployment	0.01	0.03	0.02	0.33	0.34	0.00	-0.01	0.13	-0.09	0.02	1.00		
	(0.95)	(0.79)	(0.85)	(0.00)	(0.00)	(0.99)	(0.96)	(0.26)	(0.44)	(0.87)			
S* Tone GDP	-0.36	-0.34	-0.24	0.20	0.16	-0.22	-0.31	-0.16	-0.25	-0.18	-0.03	1.00	
	(0.00)	(0.00)	(0.04)	(0.08)	(0.16)	(0.05)	(0.01)	(0.16)	(0.03)	(0.12)	(0.83)		
S* Tone Inflation	-0.06	0.07	0.01	0.15	0.22	0.06	-0.00	-0.08	-0.11	-0.11	-0.13	-0.08	1.00
	(0.63)	(0.55)	(0.92)	(0.18)	(0.06)	(0.60)	(0.99)	(0.50)	(0.34)	(0.33)	(0.26)	(0.50)	

Table A3—Unemployment: OLS Regressions

	(1) UNEM	(2) UNEM	(3) UNEM	(4) UNEM	(5) UNEM	(6) UNEM	(7) UNEM	(8) UNEM	(9) UNEM	(10) UNEM	(11) UNEM	(12) UNEM	(13) UNEM
P UNEM	0.61	0.65*	0.87**	0.65*	0.76*	0.84**	0.52	0.74*	0.51	0.66*	0.87**	0.64*	0.72***
	(1.54)	(1.67)	(2.56)	(1.75)	(1.83)	(2.51)	(1.29)	(1.80)	(1.21)	(1.69)	(2.17)	(1.84)	(2.65)
S UNEM	0.30	0.17	-0.00	0.27	0.18	0.13	0.36	0.17	0.44	0.29	0.07	0.31	0.13
	(0.78)	(0.47)	(-0.01)	(0.76)	(0.44)	(0.41)	(0.95)	(0.43)	(1.09)	(0.76)	(0.18)	(0.90)	(0.51)
P* Tone LM		0.26**											0.02
		(2.19)											(0.27)
S* Tone LM		-0.01											0.12*
		(-0.09)											(1.68)
P* Tone Bert			0.03**										
			(2.04)										
S* Tone Bert			0.03***										0.02**
			(2.64)										(2.22)
P* Tone UNEM				0.01***									0.00
				(3.15)									(1.14)
S* Tone UNEM				-0.00									
				(-0.36)									
P* Risk Bert					0.13								
					(1.23)								
S* Risk Bert					0.23								0.48
					(0.53)								(1.48)
P* Risk Uni						1.80***							
						(2.94)							
S* Risk Uni						-2.02***							
						(-3.42)							
P* Risk HHLT							1.00**						
							(2.12)						
S* Risk HHLT							-0.65***						-0.77***
							(-3.21)						(-4.45)
P* Downsid Risk								4.24					
								(1.54)					
S* Downsid Risk								-5.43					
								(-0.72)					
P* Upsid Risk									3.86				
									(0.79)				
S* Upsid Risk									12.95**				
									(2.19)				
P* Inflat Risk										10.21**			
										(2.03)			
S* Inflat Risk										-2.72			
										(-0.54)			
P* Upsid Risk Inflat											39.63**		
											(2.37)		
S* Equiti Risk Premium											-1.24		
											(-0.11)		
P* Downsid Risk Growth												46.43***	43.30***
												(3.79)	(3.38)
S* Risk Premium Corpor												8.57	24.31***
												(1.20)	(3.34)
Constant	0.55	0.09	0.94**	0.42	0.05	-0.57	-0.52	0.36	0.09	0.06	0.22	0.07	0.60*
	(1.62)	(0.23)	(2.18)	(1.32)	(0.13)	(-1.41)	(-0.80)	(1.27)	(0.35)	(0.21)	(0.77)	(0.25)	(1.87)
R-squared	0.68	0.71	0.71	0.70	0.69	0.76	0.72	0.70	0.72	0.71	0.72	0.79	0.83
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0		0.03	0.04	0.00	0.22	0.00	0.04	0.13	0.43	0.04	0.02	0.00	0.00
P*= 0 and P = 0	0.13	0.01	0.01	0.00	0.16	0.00	0.03	0.09	0.44	0.03	0.01	0.00	0.00
S* = 0		0.93	0.01	0.72	0.59	0.00	0.00	0.47	0.03	0.59	0.92	0.23	0.00
S* = 0 and S = 0	0.44	0.89	0.03	0.72	0.75	0.00	0.01	0.73	0.08	0.66	0.98	0.29	0.00

Table A4—Unemployment: 20 Percent Quantile Regressions

	(1) UNEM	(2) UNEM	(3) UNEM	(4) UNEM	(5) UNEM	(6) UNEM	(7) UNEM	(8) UNEM	(9) UNEM	(10) UNEM	(11) UNEM	(12) UNEM	(13) UNEM
q20													
P UNEM	0.92**	0.89**	0.97**	1.13***	0.81	0.70	0.71**	0.94**	0.83**	0.70*	0.87**	0.96**	1.57***
	(2.28)	(2.18)	(2.20)	(2.98)	(1.63)	(1.64)	(2.01)	(2.00)	(2.28)	(1.90)	(2.14)	(2.39)	(3.94)
S UNEM	-0.06	-0.06	-0.12	-0.24	0.03	0.10	0.11	-0.08	-0.00	0.15	-0.00	-0.12	-0.64*
	(-0.16)	(-0.15)	(-0.29)	(-0.65)	(0.08)	(0.25)	(0.35)	(-0.18)	(-0.00)	(0.43)	(-0.00)	(-0.32)	(-1.74)
P* Tone LM		0.04											-0.04
		(0.63)											(-0.51)
S* Tone LM		0.07											0.05
		(0.84)											(0.47)
P* Tone Bert			0.01										
			(0.89)										
S* Tone Bert			0.01										0.02
			(0.84)										(0.98)
P* Tone UNEM				0.00*									0.00
				(1.86)									(1.55)
S* Tone UNEM				-0.00									
				(-0.14)									
P* Risk Bert					-0.01								
					(-0.12)								
S* Risk Bert					0.32								0.38
					(1.02)								(0.98)
P* Risk Uni						0.52							
						(1.59)							
S* Risk Uni						-0.70**							
						(-2.18)							
P* Risk HHLT							0.34						
							(1.40)						
S* Risk HHLT							-0.30*						-0.30*
							(-1.73)						(-1.66)
P* Downsid Risk								-0.07					
								(-0.08)					
S* Downsid Risk								-0.05					
								(-0.01)					
P* Upsid Risk									-2.39				
									(-0.80)				
S* Upsid Risk									3.87				
									(0.83)				
P* Inflat Risk										2.18			
										(0.89)			
S* Inflat Risk										1.84			
										(0.44)			
P* Upsid Risk Inflat											8.60		
											(0.79)		
S* Equiti Risk Premium											12.13		
											(0.76)		
P* Downsid Risk Growth												19.91**	25.10**
												(2.13)	(2.35)
S* Risk Premium Corpor												18.88	19.80
												(1.45)	(1.40)
Constant	0.24	0.02	0.37	0.02	0.31	0.40	0.06	0.22	0.40	0.20	0.12	0.25	-0.02
	(0.62)	(0.04)	(1.02)	(0.05)	(0.55)	(0.81)	(0.12)	(0.60)	(0.86)	(0.58)	(0.32)	(0.68)	(-0.05)
R-squared	0.56	0.56	0.56	0.57	0.56	0.57	0.57	0.56	0.56	0.56	0.56	0.59	0.63
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0	0.02	0.53	0.38	0.06	0.90	0.11	0.16	0.94	0.42	0.38	0.43	0.04	0.04
P*= 0 and P = 0		0.10	0.09	0.01	0.07	0.09	0.04	0.14	0.04	0.10	0.09	0.01	0.00
S* = 0		0.40	0.40	0.89	0.31	0.03	0.09	0.99	0.41	0.66	0.45	0.15	0.11
S* = 0 and S = 0	0.87	0.70	0.69	0.80	0.55	0.09	0.21	0.98	0.71	0.83	0.73	0.33	0.11

Table A5—**Unemployment: 50 Percent Quantile Regressions**

	(1) UNEM	(2) UNEM	(3) UNEM	(4) UNEM	(5) UNEM	(6) UNEM	(7) UNEM	(8) UNEM	(9) UNEM	(10) UNEM	(11) UNEM	(12) UNEM	(13) UNEM
q50													
P UNEM	0.90**	0.93**	1.27**	0.95***	1.19**	0.94*	0.56	0.79*	0.88*	0.78*	1.30***	1.18***	1.04***
	(2.05)	(2.58)	(2.32)	(2.83)	(2.13)	(1.78)	(1.15)	(1.68)	(1.88)	(1.86)	(2.73)	(2.65)	(3.06)
S UNEM	0.09	-0.01	-0.29	0.03	-0.19	0.03	0.38	0.18	0.12	0.19	-0.32	-0.23	-0.15
	(0.21)	(-0.02)	(-0.55)	(0.09)	(-0.36)	(0.06)	(0.83)	(0.41)	(0.27)	(0.46)	(-0.68)	(-0.52)	(-0.44)
P* Tone LM		0.09											-0.03
		(1.58)											(-0.30)
S* Tone LM		0.04											0.16
		(0.67)											(1.65)
P* Tone Bert			0.02										
			(1.33)										
S* Tone Bert			0.00										-0.00
			(0.26)										(-0.06)
P* Tone UNEM				0.00**									0.00
				(2.33)									(0.37)
S* Tone UNEM				-0.00									
				(-1.00)									
P* Risk Bert					0.02								
					(0.33)								
S* Risk Bert					0.24								0.35
					(0.61)								(0.84)
P* Risk Uni						0.88*							
						(1.80)							
S* Risk Uni						-0.96**							
						(-2.26)							
P* Risk Hassan							0.43*						
							(1.82)						
S* Risk Hassan							-0.30**						-0.48**
							(-2.53)						(-2.26)
P* Downsid Risk								0.72					
								(0.51)					
S* Downsid Risk								-5.34*					
								(-1.74)					
P* Upsid Risk									0.95				
									(0.26)				
S* Upsid Risk									6.93				
									(0.79)				
P* Inflat Risk										2.37			
										(0.87)			
S* Inflat Risk										-3.19			
										(-0.97)			
P* Upsid Risk Inflat											15.92		
											(1.14)		
S* Equiti Risk Premium											3.39		
											(0.29)		
P* Downsid Risk Growth												27.92**	30.06*
												(2.03)	(1.89)
S* Risk Premium Corpor												9.07	15.18
												(1.12)	(0.73)
Constant	-0.11	-0.25	-0.00	-0.08	-0.19	-0.30	-0.33	0.01	-0.22	-0.08	-0.14	0.03	0.20
	(-0.49)	(-0.80)	(-0.00)	(-0.39)	(-0.66)	(-0.87)	(-0.84)	(0.04)	(-0.81)	(-0.36)	(-0.58)	(0.11)	(0.38)
R-squared	0.57	0.58	0.57	0.58	0.57	0.59	0.58	0.57	0.57	0.58	0.58	0.61	0.64
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0	0.04	0.12	0.18	0.02	0.74	0.07	0.07	0.61	0.80	0.38	0.26	0.04	0.30
P*= 0 and P = 0		0.00	0.07	0.00	0.07	0.14	0.09	0.22	0.14	0.14	0.02	0.01	0.02
S* = 0		0.51	0.79	0.32	0.54	0.03	0.01	0.08	0.43	0.33	0.77	0.27	0.22
S* = 0 and S = 0	0.84	0.79	0.82	0.59	0.82	0.03	0.03	0.22	0.73	0.59	0.79	0.47	0.31

Table A6—Unemployment: 80 Percent Quantile Regressions

	(1) UNEM	(2) UNEM	(3) UNEM	(4) UNEM	(5) UNEM	(6) UNEM	(7) UNEM	(8) UNEM	(9) UNEM	(10) UNEM	(11) UNEM	(12) UNEM	(13) UNEM
q80													
P UNEM	0.94*	0.39	0.85*	0.44	0.86*	0.72	0.87	0.88**	0.64	0.86	0.90	0.49	0.64**
	(1.84)	(0.75)	(1.86)	(0.84)	(1.70)	(1.40)	(1.49)	(2.23)	(0.94)	(1.46)	(1.61)	(0.84)	(2.17)
S UNEM	0.00	0.45	0.08	0.44	0.09	0.37	0.05	0.11	0.39	0.22	0.12	0.55	0.30
	(0.01)	(0.91)	(0.17)	(0.79)	(0.19)	(0.68)	(0.09)	(0.27)	(0.59)	(0.38)	(0.22)	(0.94)	(0.92)
P* Tone LM		0.27											-0.07
		(1.18)											(-0.74)
S* Tone LM		0.02											0.24**
		(0.13)											(2.47)
P* Tone Bert			0.01										
			(0.55)										
S* Tone Bert			0.03*										0.02
			(1.84)										(1.14)
P* Tone UNEM				0.01*									0.01*
				(1.72)									(1.68)
S* Tone UNEM				-0.00									
				(-1.01)									
P* Risk Bert					0.03								
					(0.17)								
S* Risk Bert					0.13								0.82**
					(0.17)								(2.05)
P* Risk Uni						2.05*							
						(1.97)							
S* Risk Uni						-2.32**							
						(-2.37)							
P* Risk Hassan							0.66						
							(0.86)						
S* Risk Hassan							-0.77*						-0.90***
							(-1.83)						(-4.15)
P* Downsid Risk								7.01					
								(1.39)					
S* Downsid Risk								-13.77					
								(-1.29)					
P* Upsid Risk									1.26				
									(0.13)				
S* Upsid Risk									24.54*				
									(1.88)				
P* Inflat Risk										16.67			
										(1.47)			
S* Inflat Risk										-10.34			
										(-0.51)			
P* Upsid Risk Inflat											59.31*		
											(1.67)		
S* Equiti Risk Premium											12.07		
											(0.56)		
P* Downsid Risk Growth												70.94***	70.76***
												(3.41)	(3.37)
S* Risk Premium Corpor												9.62	16.15
												(0.64)	(0.90)
Constant	0.65	0.24	0.92	0.97	0.48	-0.80	0.22	0.21	-0.00	-0.40	-0.03	-0.11	0.44
	(0.89)	(0.35)	(1.22)	(1.06)	(0.48)	(-1.02)	(0.19)	(0.27)	(-0.00)	(-0.41)	(-0.03)	(-0.20)	(0.63)
R-squared	0.47	0.49	0.51	0.49	0.47	0.53	0.50	0.50	0.52	0.48	0.51	0.58	0.66
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0	0.07	0.24	0.59	0.09	0.86	0.05	0.39	0.17	0.90	0.14	0.10	0.00	0.00
P*= 0 and P = 0		0.38	0.17	0.18	0.20	0.08	0.22	0.06	0.57	0.09	0.06	0.00	0.00
S* = 0		0.90	0.07	0.31	0.86	0.02	0.07	0.20	0.06	0.61	0.57	0.52	0.00
S* = 0 and S = 0	0.99	0.66	0.19	0.49	0.97	0.04	0.19	0.43	0.18	0.81	0.83	0.59	0.00

Table A7—Real GDP Growth: OLS Regressions

	(1) Real GDP	(2) Real GDP	(3) Real GDP	(4) Real GDP	(5) Real GDP	(6) Real GDP	(7) Real GDP	(8) Real GDP	(9) Real GDP	(10) Real GDP	(11) Real GDP	(12) Real GDP	(13) Real GDP
P GDP	0.54	0.60	0.56	0.55	0.59*	0.63*	0.64*	0.61*	0.54	0.64*	0.69**	0.65**	0.66***
	(1.53)	(1.59)	(1.58)	(1.54)	(1.72)	(1.95)	(1.81)	(1.76)	(1.64)	(1.87)	(1.98)	(2.08)	(5.85)
S GDP	0.22	0.12	0.12	0.22	0.19	0.17	0.13	0.15	0.22	0.15	0.08	0.10	
	(0.76)	(0.34)	(0.40)	(0.72)	(0.65)	(0.59)	(0.44)	(0.49)	(0.78)	(0.51)	(0.27)	(0.37)	
P* Tone LM		-0.05											
		(-0.35)											
S* Tone LM		-0.12											-0.33***
		(-0.81)											(-3.52)
P* Tone Bert			-0.03										
			(-1.47)										
S* Tone Bert			-0.02										
			(-1.40)										
P* Tone GDP				-0.00									
				(-0.01)									
S* Tone GDP				0.00									
				(0.24)									
P* Risk Bert					-0.18**								
					(-2.03)								
S* Risk Bert					0.14								
					(0.30)								
P* Risk Uni						-1.50***							
						(-3.82)							
S* Risk Uni						0.79							
						(1.46)							
P* Risk HHLT							-0.76*						
							(-1.70)						
S* Risk HHLT							0.01						
							(0.04)						
P* Downsid Risk								-3.83*					
								(-1.85)					
S* Downsid Risk								-6.06					
								(-0.86)					
P* Upsid Risk									-8.61				
									(-1.34)				
S* Upsid Risk									-3.96				
									(-0.72)				
P* Inflat Risk										-9.18*			
										(-1.97)			
S* Inflat Risk										-10.43			-9.15
										(-1.60)			(-1.50)
P* Upsid Risk Inflat											-57.20***		-38.12*
											(-4.08)		(-1.93)
S* Equiti Risk Premium											-0.00		
											(-0.00)		
P* Downsid Risk Growth												-41.62***	-27.58***
												(-5.99)	(-2.76)
S* Risk Premium Corpor												-10.73	-12.47
												(-0.89)	(-1.13)
Constant	0.43	1.16	0.52	0.44	0.70*	1.09***	1.63**	0.75**	0.64*	0.61*	0.63*	0.67**	2.22***
	(1.25)	(1.60)	(1.48)	(1.25)	(1.89)	(2.71)	(2.12)	(2.11)	(1.89)	(1.79)	(1.86)	(2.07)	(4.29)
R-squared	0.36	0.37	0.39	0.36	0.38	0.43	0.39	0.39	0.39	0.40	0.45	0.46	0.52
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0		0.73	0.14	0.99	0.04	0.00	0.09	0.07	0.18	0.05	0.00	0.00	0.00
P*= 0 and P = 0	0.13	0.29	0.16	0.30	0.03	0.00	0.08	0.08	0.10	0.04	0.00	0.00	0.00
S* = 0		0.42	0.16	0.81	0.77	0.15	0.97	0.39	0.47	0.11	1.00	0.38	0.00
S* = 0 and S = 0	0.45	0.63	0.35	0.71	0.80	0.30	0.90	0.66	0.58	0.24	0.95	0.48	0.00

Table A8—Real GDP Growth: 20 Percent Quantile Regressions

	(1) Real GDP	(2) Real GDP	(3) Real GDP	(4) Real GDP	(5) Real GDP	(6) Real GDP	(7) Real GDP	(8) Real GDP	(9) Real GDP	(10) Real GDP	(11) Real GDP	(12) Real GDP	(13) Real GDP
q20													
P GDP	0.63	0.91*	0.70	0.68	0.71	1.00*	0.92*	0.84	0.82	0.85	0.61	0.71	0.66***
	(1.06)	(1.67)	(1.15)	(1.06)	(1.45)	(1.93)	(1.77)	(1.53)	(1.31)	(1.49)	(1.23)	(1.42)	(3.08)
S GDP	0.41	-0.17	-0.01	0.41	0.40	-0.10	0.22	0.21	0.23	0.22	0.40	0.29	
	(0.75)	(-0.33)	(-0.02)	(0.68)	(0.80)	(-0.20)	(0.44)	(0.38)	(0.40)	(0.39)	(0.87)	(0.57)	
P* Tone LM		-0.35***											
		(-2.88)											
S* Tone LM		-0.12											-0.59***
		(-0.54)											(-3.24)
P* Tone Bert			-0.06*										
			(-1.84)										
S* Tone Bert			-0.03										
			(-1.29)										
P* Tone GDP				0.18									
				(0.32)									
S* Tone GDP				0.01									
				(0.03)									
P* Risk Bert					-12.21								
					(-0.67)								
S* Risk Bert					14.29								
					(0.20)								
P* Risk Uni						-1.17**							
						(-2.22)							
S* Risk Uni						1.65**							
						(1.98)							
P* Risk HHLT							-1.07						
							(-1.65)						
S* Risk HHLT							0.44						
							(0.99)						
P* Downsid Risk								-5.05					
								(-1.51)					
S* Downsid Risk								-4.81					
								(-0.37)					
P* Upsid Risk									-19.92*				
									(-1.69)				
S* Upsid Risk									1.52				
									(0.13)				
P* Inflat Risk										-12.40**			
										(-2.16)			
S* Inflat Risk										-3.52			-0.56
										(-0.33)			(-0.05)
P* Upsid Risk Inflat											-59.88***		-83.51**
											(-3.37)		(-2.08)
S* Equiti Risk Premium											35.54		
											(1.46)		
P* Downsid Risk Growth												-24.00*	-10.78
												(-1.68)	(-0.52)
S* Risk Premium Corpor												9.99	-6.98
												(0.33)	(-0.24)
Constant	-1.28**	1.32	-0.56	-1.35**	-1.35**	-0.67	-0.11	-1.07*	-1.13*	-1.16**	-1.15***	-1.11**	2.27*
	(-2.21)	(0.99)	(-0.91)	(-2.09)	(-2.36)	(-1.21)	(-0.11)	(-1.72)	(-1.79)	(-2.30)	(-2.66)	(-2.05)	(1.89)
R-squared	0.25	0.29	0.30	0.25	0.26	0.30	0.27	0.27	0.30	0.28	0.33	0.31	0.38
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0		0.00	0.07	0.75	0.50	0.03	0.10	0.13	0.09	0.03	0.00	0.10	0.00
P*= 0 and P = 0	0.29	0.00	0.15	0.45	0.31	0.03	0.07	0.16	0.18	0.02	0.00	0.12	0.00
S* = 0		0.59	0.20	0.98	0.84	0.05	0.32	0.72	0.90	0.74	0.15	0.74	0.01
S* = 0 and S = 0	0.46	0.84	0.42	0.79	0.67	0.14	0.58	0.90	0.91	0.90	0.31	0.81	0.01

Table A9—Real GDP Growth: 50 Percent Quantile Regressions

	(1) Real GDP	(2) Real GDP	(3) Real GDP	(4) Real GDP	(5) Real GDP	(6) Real GDP	(7) Real GDP	(8) Real GDP	(9) Real GDP	(10) Real GDP	(11) Real GDP	(12) Real GDP	(13) Real GDP
q50													
P GDP	0.38	0.54	0.18	0.64	0.53	0.42	0.71	0.51	0.72	0.78*	0.91**	0.69	0.49***
	(0.73)	(0.94)	(0.31)	(1.17)	(1.23)	(0.73)	(1.25)	(0.96)	(1.53)	(1.66)	(2.00)	(1.65)	(3.46)
S GDP	0.27	0.06	0.35	0.05	0.20	0.37	-0.04	0.14	-0.05	-0.08	-0.18	-0.04	
	(0.64)	(0.13)	(0.77)	(0.11)	(0.62)	(0.75)	(-0.08)	(0.33)	(-0.13)	(-0.22)	(-0.44)	(-0.12)	
P* Tone LM		-0.03											
		(-0.18)											
S* Tone LM		-0.24											-0.38**
		(-1.44)											(-2.40)
P* Tone Bert			-0.01										
			(-0.62)										
S* Tone Bert			-0.03										
			(-1.10)										
P* Tone GDP				0.04									
				(0.12)									
S* Tone GDP				0.18									
				(0.87)									
P* Risk Bert					-14.19								
					(-1.27)								
S* Risk Bert					11.77								
					(0.16)								
P* Risk Uni						-1.02							
						(-1.52)							
S* Risk Uni						0.27							
						(0.31)							
P* Risk HHLT							-0.57						
							(-1.16)						
S* Risk HHLT							-0.06						
							(-0.22)						
P* Downsid Risk								-3.26					
								(-1.01)					
S* Downsid Risk								0.41					
								(0.04)					
P* Upsid Risk									-3.58				
									(-0.39)				
S* Upsid Risk									-13.10				
									(-1.01)				
P* Inflat Risk										-4.09			
										(-0.58)			
S* Inflat Risk										-12.38			-14.34
										(-1.59)			(-1.20)
P* Upsid Risk Inflat											-66.93***		-11.66
											(-2.62)		(-0.41)
S* Equiti Risk Premium											-9.09		
											(-0.47)		
P* Downsid Risk Growth												-43.15***	-38.42***
												(-5.82)	(-3.10)
S* Risk Premium Corpor												-10.56	-17.29
												(-0.62)	(-0.92)
Constant	0.84*	1.84**	0.98**	0.72*	1.01*	1.02*	1.68**	1.03**	0.87*	0.87*	0.84**	1.00***	2.81***
	(1.81)	(2.13)	(2.05)	(1.69)	(1.72)	(1.83)	(2.42)	(2.06)	(1.87)	(1.92)	(2.01)	(2.99)	(4.12)
R-squared	0.20	0.22	0.22	0.20	0.21	0.22	0.21	0.21	0.21	0.22	0.24	0.27	0.31
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0		0.86	0.54	0.90	0.21	0.13	0.25	0.31	0.69	0.56	0.01	0.00	0.00
P*= 0 and P = 0	0.47	0.65	0.78	0.49	0.20	0.29	0.39	0.47	0.30	0.24	0.02	0.00	0.00
S* = 0		0.15	0.27	0.38	0.87	0.76	0.83	0.97	0.31	0.11	0.64	0.54	0.06
S* = 0 and S = 0	0.53	0.35	0.47	0.59	0.82	0.75	0.96	0.94	0.56	0.28	0.86	0.82	0.06

Table A10—**Real GDP Growth: 80 Percent Quantile Regressions**

	(1) Real GDP	(2) Real GDP	(3) Real GDP	(4) Real GDP	(5) Real GDP	(6) Real GDP	(7) Real GDP	(8) Real GDP	(9) Real GDP	(10) Real GDP	(11) Real GDP	(12) Real GDP	(13) Real GDP
q80													
P GDP	0.13	-0.02	0.24	0.10	0.11	0.61	-0.18	0.22	-0.04	0.70	0.58	0.64	0.58***
	(0.18)	(-0.02)	(0.41)	(0.15)	(0.21)	(1.07)	(-0.34)	(0.31)	(-0.05)	(1.00)	(0.87)	(1.04)	(4.66)
S GDP	0.50	0.59	0.36	0.53	0.58	0.12	0.81*	0.47	0.71	0.00	0.11	0.08	
	(0.84)	(0.89)	(0.75)	(0.96)	(1.30)	(0.25)	(1.70)	(0.74)	(1.00)	(0.00)	(0.20)	(0.14)	
P* Tone LM		0.31											
		(1.25)											
S* Tone LM		-0.32											-0.23
		(-1.23)											(-1.28)
P* Tone Bert			-0.04										
			(-1.55)										
S* Tone Bert			0.02										
			(0.46)										
P* Tone GDP				-0.01									
				(-0.01)									
S* Tone GDP				0.21									
				(0.78)									
P* Risk Bert					-38.58***								
					(-4.52)								
S* Risk Bert					-20.97								
					(-0.29)								
P* Risk Uni						-2.15***							
						(-4.12)							
S* Risk Uni						0.49							
						(0.72)							
P* Risk HHLT							-0.60						
							(-1.05)						
S* Risk HHLT							-0.42						
							(-1.11)						
P* Downsid Risk								-5.17					
								(-1.61)					
S* Downsid Risk								-8.05					
								(-0.82)					
P* Upsid Risk									-6.58				
									(-1.08)				
S* Upsid Risk									6.89				
									(0.61)				
P* Inflat Risk										-5.35			
										(-0.99)			
S* Inflat Risk										-28.37***			-23.92**
										(-3.28)			(-2.10)
P* Upsid Risk Inflat											-62.27***		-24.03
											(-2.77)		(-1.38)
S* Equiti Risk Premium											-30.82*		
											(-1.74)		
P* Downsid Risk Growth												-51.37***	-39.80**
												(-3.64)	(-2.42)
S* Risk Premium Corpor												-35.96	-16.16
												(-1.55)	(-1.11)
Constant	1.77***	1.82*	1.77***	1.75***	2.56***	2.64***	3.07***	2.03***	1.82***	1.83***	2.00***	1.87***	3.02***
	(3.55)	(1.69)	(4.12)	(3.83)	(7.43)	(6.26)	(3.59)	(4.62)	(2.96)	(3.32)	(4.55)	(4.45)	(4.02)
R-squared	0.20	0.21	0.21	0.21	0.26	0.30	0.25	0.23	0.21	0.24	0.28	0.28	0.33
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0		0.21	0.12	0.99	0.00	0.00	0.30	0.11	0.28	0.33	0.01	0.00	0.00
P*= 0 and P = 0	0.86	0.44	0.29	0.99	0.00	0.00	0.55	0.28	0.52	0.54	0.01	0.00	0.00
S* = 0		0.22	0.64	0.44	0.78	0.47	0.27	0.41	0.54	0.00	0.08	0.12	0.02
S* = 0 and S = 0	0.40	0.32	0.66	0.41	0.40	0.72	0.12	0.54	0.46	0.01	0.06	0.19	0.02

Table A11—**Inflation: OLS Regressions**

	(1) INFL	(2) INFL	(3) INFL	(4) INFL	(5) INFL	(6) INFL	(7) INFL	(8) INFL	(9) INFL	(10) INFL	(11) INFL	(12) INFL	(13) INFL
P INFL	0.25	0.34	0.26	0.29	0.28	0.31	0.25	0.25	0.22	0.29	0.25	0.25	0.53*
	(0.87)	(1.17)	(0.84)	(0.98)	(1.01)	(0.98)	(0.84)	(0.86)	(0.76)	(0.97)	(0.87)	(0.83)	(1.77)
S INFL	0.65**	0.59**	0.65**	0.60**	0.68**	0.64**	0.65**	0.66**	0.69**	0.61**	0.67**	0.65**	0.49
	(2.21)	(2.05)	(2.09)	(2.05)	(2.44)	(1.99)	(2.09)	(2.24)	(2.39)	(2.01)	(2.30)	(2.09)	(1.65)
P* Tone LM		-0.10											-0.13*
		(-1.46)											(-1.74)
S* Tone LM		-0.03											-0.07
		(-0.33)											(-0.67)
P* Tone Bert			-0.01										
			(-1.36)										
S* Tone Bert			0.00										
			(0.39)										
P* Tone Inflation				-0.04									
				(-0.13)									
S* Tone Inflation				0.16									0.21
				(0.98)									(1.16)
P* Risk Bert					13.43**								7.34
					(2.52)								(1.00)
S* Risk Bert					6.93								
					(0.28)								
P* Risk Uni						0.46*							0.91**
						(1.74)							(2.33)
S* Risk Uni						-0.01							
						(-0.03)							
P* Risk HHLT							-0.06						-0.47**
							(-0.29)						(-2.00)
S* Risk HHLT							0.05						-0.20
							(0.25)						(-1.01)
P* Downsid Risk								0.46					
								(0.55)					
S* Downsid Risk								1.13					
								(0.47)					
P* Upsid Risk									3.74**				
									(2.07)				
S* Upsid Risk									-4.51***				-5.53***
									(-2.64)				(-2.76)
P* Inflat Risk										1.62			
										(1.32)			
S* Inflat Risk										-1.55			
										(-0.51)			
P* Upsid Risk Inflat											5.42		
											(1.43)		
S* Equiti Risk Premium											9.96		13.55**
											(1.56)		(2.20)
P* Downsid Risk Growth												0.21	
												(0.08)	
S* Risk Premium Corpor												0.06	
												(0.01)	
Constant	0.25**	0.66**	0.22**	0.27**	-0.20	-0.16	0.30	0.18	0.18	0.19	0.13	0.24*	0.79**
	(2.35)	(2.53)	(2.03)	(2.42)	(-0.92)	(-0.71)	(0.91)	(1.16)	(1.39)	(1.48)	(0.84)	(1.84)	(2.24)
R-squared	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.89
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0		0.15	0.18	0.90	0.01	0.08	0.77	0.58	0.04	0.19	0.16	0.94	0.00
P*= 0 and P = 0	0.38	0.09	0.24	0.62	0.04	0.22	0.51	0.63	0.11	0.33	0.30	0.66	0.00
S* = 0		0.75	0.70	0.33	0.78	0.97	0.80	0.64	0.01	0.61	0.12	0.99	0.02
S* = 0 and S = 0	0.03	0.05	0.09	0.08	0.05	0.05	0.05	0.08	0.01	0.06	0.04	0.10	0.00

Table A12—**Inflation: 20 Percent Quantile Regressions**

	(1) INFL	(2) INFL	(3) INFL	(4) INFL	(5) INFL	(6) INFL	(7) INFL	(8) INFL	(9) INFL	(10) INFL	(11) INFL	(12) INFL	(13) INFL
q20													
P INFL	0.35	0.22	0.30	0.32	0.68	0.57	0.39	0.38	0.31	0.35	0.47	0.41	0.66
	(0.90)	(0.47)	(0.69)	(0.79)	(1.52)	(1.09)	(0.85)	(0.78)	(0.65)	(0.70)	(1.01)	(0.91)	(1.55)
S INFL	0.46	0.53	0.44	0.40	0.19	0.28	0.45	0.43	0.52	0.47	0.36	0.36	0.27
	(1.26)	(1.27)	(1.15)	(1.18)	(0.51)	(0.61)	(1.08)	(1.04)	(1.28)	(1.03)	(0.90)	(0.92)	(0.71)
P* Tone LM		-0.06											-0.08
		(-0.78)											(-0.90)
S* Tone LM		-0.06											-0.06
		(-0.51)											(-0.52)
P* Tone Bert			-0.00										
			(-0.46)										
S* Tone Bert			-0.01										
			(-0.63)										
P* Tone Inflation				-0.29									
				(-1.14)									
S* Tone Inflation				0.16									0.22
				(0.90)									(0.88)
P* Risk Bert					10.41								0.12
					(1.55)								(1.34)
S* Risk Bert					30.85								
					(1.17)								
P* Risk Uni						0.04							0.59
						(0.14)							(1.14)
S* Risk Uni						0.37							
						(0.87)							
P* Risk HHLT							0.02						-0.45
							(0.11)						(-1.14)
S* Risk HHLT							0.16						0.13
							(0.64)						(0.57)
P* Downsid Risk								0.11					
								(0.11)					
S* Downsid Risk								3.54					
								(1.25)					
P* Upsid Risk									2.98				
									(1.50)				
S* Upsid Risk									-2.65				-2.76
									(-0.78)				(-0.89)
P* Inflat Risk										0.02			
										(0.01)			
S* Inflat Risk										1.52			
										(0.29)			
P* Upsid Risk Inflat											12.25		
											(1.31)		
S* Equiti Risk Premium											1.02		4.67
											(0.12)		(0.53)
P* Downsid Risk Growth												-1.42	
												(-0.31)	
S* Risk Premium Corpor												-9.99	
												(-0.92)	
Constant	0.07	0.62*	0.19	0.26	-0.38	-0.14	-0.14	0.05	0.03	0.06	-0.02	0.26	0.22
	(0.28)	(1.87)	(0.77)	(0.98)	(-0.84)	(-0.30)	(-0.25)	(0.16)	(0.11)	(0.17)	(-0.05)	(0.80)	(0.41)
R-squared	0.43	0.44	0.44	0.45	0.44	0.44	0.44	0.44	0.45	0.43	0.44	0.44	0.48
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0		0.44	0.64	0.26	0.12	0.89	0.91	0.92	0.14	0.99	0.19	0.76	0.19
P*= 0 and P = 0	0.37	0.56	0.59	0.36	0.19	0.52	0.67	0.73	0.29	0.77	0.34	0.52	0.24
S* = 0		0.61	0.53	0.37	0.24	0.38	0.52	0.21	0.44	0.77	0.90	0.36	0.70
S* = 0 and S = 0	0.21	0.21	0.28	0.29	0.30	0.61	0.51	0.30	0.18	0.58	0.65	0.46	0.76

Table A13—**Inflation: 50 Percent Quantile Regressions**

	(1) INFL	(2) INFL	(3) INFL	(4) INFL	(5) INFL	(6) INFL	(7) INFL	(8) INFL	(9) INFL	(10) INFL	(11) INFL	(12) INFL	(13) INFL
q50													
P INFL	0.25	0.43	0.43	0.28	0.34	0.33	0.29	0.15	0.31	0.42	0.41	0.24	0.43
	(0.70)	(1.18)	(1.26)	(1.08)	(1.30)	(0.92)	(0.90)	(0.50)	(1.03)	(1.37)	(1.46)	(0.83)	(1.41)
P INFL	0.64*	0.48	0.46	0.60**	0.60**	0.57	0.58	0.75**	0.58*	0.46	0.46	0.64**	0.60*
	(1.73)	(1.26)	(1.22)	(2.15)	(2.26)	(1.48)	(1.57)	(2.40)	(1.86)	(1.39)	(1.65)	(2.13)	(1.88)
P* Tone LM		-0.08											-0.06
		(-1.46)											(-0.99)
S* Tone LM		-0.06											-0.13
		(-0.62)											(-1.53)
P* Tone Bert			-0.01										
			(-0.72)										
S* Tone Bert			-0.00										
			(-0.26)										
S* Tone Inflation				-0.16									
				(-0.51)									
P* Tone Inflation				-0.15									0.05
				(-0.75)									(0.38)
P* Risk Bert					13.72**								14.28*
					(2.12)								(1.72)
S* Risk Bert					-14.94								
					(-0.46)								
P* Risk Uni						0.28							0.19
						(0.93)							(0.58)
S* Risk Uni						0.09							
						(0.22)							
P* Risk HHLT							0.05						-0.09
							(0.32)						(-0.38)
S* Risk HHLT							-0.06						-0.03
							(-0.32)						(-0.19)
P* Downsid Risk								0.23					
								(0.23)					
S* Downsid Risk								5.13					
								(1.51)					
P* Upsid Risk									2.78				
									(1.48)				
S* Upsid Risk									-4.47				-6.60**
									(-1.47)				(-2.31)
P* Inflat Risk										1.81			
										(0.97)			
S* Inflat Risk										-0.35			
										(-0.10)			
P* Upsid Risk Inflat											8.55*		
											(1.66)		
S* Equiti Risk Premium											3.54		7.87
											(0.41)		(0.91)
P* Downsid Risk Growth												0.81	
												(0.22)	
S* Risk Premium Corpor												-0.97	
												(-0.11)	
Constant	0.27**	0.73***	0.25*	0.28**	-0.19	0.01	0.26	0.11	0.22*	0.19	0.18	0.27*	0.34
	(2.15)	(2.99)	(1.70)	(2.17)	(-0.73)	(0.03)	(0.72)	(0.62)	(1.77)	(1.19)	(0.87)	(1.77)	(0.85)
R-squared	0.58	0.60	0.59	0.59	0.60	0.59	0.58	0.59	0.59	0.59	0.59	0.58	0.63
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0		0.15	0.47	0.61	0.04	0.35	0.75	0.82	0.14	0.34	0.10	0.82	0.02
P*= 0 and P = 0	0.49	0.08	0.37	0.46	0.10	0.49	0.62	0.87	0.24	0.34	0.19	0.70	0.03
S* = 0		0.53	0.80	0.45	0.65	0.82	0.75	0.13	0.14	0.92	0.68	0.91	0.11
S* = 0 and S = 0	0.09	0.09	0.19	0.06	0.05	0.31	0.14	0.04	0.13	0.34	0.20	0.09	0.04

Table A14—**Inflation: 80 Percent Quantile Regressions**

	(1) INFL	(2) INFL	(3) INFL	(4) INFL	(5) INFL	(6) INFL	(7) INFL	(8) INFL	(9) INFL	(10) INFL	(11) INFL	(12) INFL	(13) INFL
q80													
P INFL	0.72	0.44	0.61	0.52	0.26	0.77	0.48	0.39	0.64	0.76	0.38	0.62	1.03*
	(1.61)	(1.00)	(1.17)	(1.15)	(0.62)	(1.42)	(0.89)	(0.93)	(1.50)	(1.43)	(0.81)	(1.37)	(1.89)
S INFL	0.26	0.59	0.39	0.50	0.77*	0.26	0.48	0.58	0.35	0.23	0.62	0.36	0.11
	(0.57)	(1.30)	(0.74)	(1.05)	(1.72)	(0.47)	(0.83)	(1.33)	(0.77)	(0.42)	(1.26)	(0.75)	(0.20)
P* Tone LM		-0.04											-0.06
		(-0.49)											(-0.66)
S* Tone LM		-0.10											-0.17
		(-0.80)											(-1.13)
P* Tone Bert			0.00										
			(0.02)										
S* Tone Bert			-0.01										
			(-0.61)										
P* Tone Inflation				-0.35									
				(-0.77)									
S* Tone Inflation				-0.15									0.09
				(-0.69)									(0.34)
P* Risk Bert					12.64								-2.73
					(1.41)								(-0.19)
S* Risk Bert					8.56								
					(0.19)								
P* Risk Uni						0.97							1.89**
						(1.64)							(2.03)
S* Risk Uni						-0.68							
						(-1.24)							
P* Risk HHLT							-0.26						-0.56*
							(-0.75)						(-1.85)
S* Risk HHLT							-0.12						-0.35
							(-0.52)						(-0.98)
P* Downsid Risk								-0.31					
								(-0.16)					
S* Downsid Risk								-4.26					
								(-0.95)					
P* Upsid Risk									-0.54				
									(-0.12)				
S* Upsid Risk									-2.27				-9.17*
									(-0.56)				(-1.71)
P* Inflat Risk										0.35			
										(0.11)			
S* Inflat Risk										-4.38			
										(-1.30)			
P* Upsid Risk Inflat											-1.38		
											(-0.24)		
S* Equiti Risk Premium											12.76		13.01
											(1.05)		(0.98)
P* Downsid Risk Growth												-2.83	
												(-0.54)	
S* Risk Premium Corpor												-3.11	
												(-0.38)	
Constant	0.39***	0.73**	0.30*	0.27*	-0.03	-0.19	0.95**	0.52**	0.43**	0.38**	0.35**	0.43***	0.79
	(2.67)	(2.34)	(1.91)	(1.82)	(-0.08)	(-0.48)	(1.99)	(2.42)	(2.56)	(2.21)	(2.10)	(3.42)	(0.37)
R-squared	0.69	0.70	0.69	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.69	0.73
Observations	129	129	129	129	129	129	129	129	129	129	129	129	129
P* = 0		0.63	0.98	0.44	0.16	0.10	0.45	0.87	0.90	0.91	0.81	0.59	0.02
P*= 0 and P = 0	0.11	0.52	0.50	0.35	0.30	0.19	0.35	0.65	0.28	0.32	0.68	0.13	0.04
S* = 0		0.42	0.54	0.49	0.85	0.22	0.60	0.34	0.58	0.20	0.30	0.71	0.24
S* = 0 and S = 0	0.57	0.23	0.50	0.47	0.20	0.18	0.30	0.29	0.68	0.32	0.34	0.71	0.08

Table A15—**Unemployment Regressions: Romer and Romer Sample: 1979-2001**

	(1) UNEM	(2) UNEM	(3) UNEM	(4) UNEM	(5) UNEM	(6) UNEM	(7) UNEM	(8) UNEM	(9) UNEM	(10) UNEM	(11) UNEM	(12) UNEM	(13) UNEM
P UNEM	-0.10	0.17	0.34	-0.05	-0.21	-0.13	-0.12	-0.22	-0.15	-0.03	-0.09	-0.10	0.27
	(-0.23)	(0.37)	(0.74)	(-0.12)	(-0.45)	(-0.28)	(-0.27)	(-0.45)	(-0.30)	(-0.07)	(-0.21)	(-0.22)	(0.57)
P UNEM	1.03**	0.70	0.58	0.99**	1.09**	1.05**	1.03**	1.12**	1.03**	0.91*	1.01**	1.02**	0.60
	(2.33)	(1.58)	(1.26)	(2.37)	(2.39)	(2.38)	(2.34)	(2.39)	(2.17)	(1.94)	(2.21)	(2.30)	(1.32)
P* Tone LM		0.04											
		(0.33)											
S* Tone LM		0.19**											0.11
		(2.06)											(1.55)
P* Tone Bert			0.00										
			(0.44)										
S* Tone Bert			0.03**										0.02*
			(2.44)										(1.86)
P* Tone UNEM				0.00									
				(1.63)									
S* Tone UNEM				0.00									
				(0.62)									
P* Risk Bert					-0.12								
					(-1.11)								
S* Risk Bert					-0.17								
					(-0.41)								
P* Risk Uni						-0.18							
						(-0.28)							
S* Risk Uni						0.59							
						(0.58)							
P* Risk HHLT							0.13						
							(0.43)						
S* Risk HHLT							-0.10						
							(-0.18)						
P* Downsid Risk								-1.28					
								(-1.09)					
S* Downsid Risk								-6.84*					
								(-1.72)					
P* Upsid Risk									-7.27				-7.25
									(-1.45)				(-1.59)
S* Upsid Risk									-1.41				
									(-0.37)				
P* Inflat Risk										-8.12*			
										(-1.72)			
S* Inflat Risk										5.73			
										(0.90)			
P* Upsid Risk Inflat											-8.26		
											(-0.37)		
S* Equiti Risk Premium											-6.18		
											(-0.65)		
P* Downsid Risk Growth												-18.62**	
												(-2.48)	
P* Risk Premium Corpor												0.00	
												(.)	
Constant	0.33	-0.25	0.46	0.22	0.78	0.34	0.26	0.53	0.66	0.71	0.37	0.34	0.40
	(0.92)	(-0.65)	(1.24)	(0.67)	(1.32)	(0.45)	(0.30)	(1.35)	(1.56)	(1.46)	(0.88)	(0.93)	(0.87)
R-squared	0.78	0.80	0.80	0.79	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.81
Observations	69	69	69	69	69	69	69	69	69	69	69	69	69
P* = 0		0.74	0.66	0.11	0.27	0.78	0.67	0.28	0.15	0.09	0.71	0.02	0.12
P*= 0 and P = 0	0.82	0.80	0.74	0.25	0.54	0.94	0.90	0.54	0.31	0.17	0.88	0.05	0.08
S* = 0		0.04	0.02	0.54	0.68	0.56	0.86	0.09	0.71	0.37	0.52	.	0.04
S* = 0 and S = 0	0.02	0.00	0.01	0.05	0.06	0.05	0.07	0.04	0.02	0.16	0.03	0.02	0.00

Table A16—Real GDP Growth Regressions: Romer and Romer Sample: 1979-2001

	(1) Real GDP	(2) Real GDP	(3) Real GDP	(4) Real GDP	(5) Real GDP	(6) Real GDP	(7) Real GDP	(8) Real GDP	(9) Real GDP	(10) Real GDP	(11) Real GDP	(12) Real GDP	(13) Real GDP
P GDP	0.63	0.66	0.70	0.52	0.64	0.70	0.66	0.63	0.68	0.80	0.63	0.61	0.66
	(1.32)	(1.17)	(1.38)	(1.04)	(1.31)	(1.40)	(1.28)	(1.31)	(1.41)	(1.60)	(1.33)	(1.28)	(1.20)
S GDP	0.25	0.11	0.07	0.32	0.22	0.19	0.22	0.26	0.21	0.05	0.24	0.28	0.09
	(0.56)	(0.20)	(0.14)	(0.71)	(0.48)	(0.40)	(0.46)	(0.57)	(0.47)	(0.10)	(0.52)	(0.61)	(0.17)
P* Tone LM		0.09											
		(0.42)											
S* Tone LM		-0.50***											-0.43***
		(-3.24)											(-3.08)
P* Tone Bert			-0.05**										
			(-2.35)										
S* Tone Bert			-0.03*										
			(-1.68)										
P* Tone GDP				-0.01*									-0.01
				(-1.69)									(-1.18)
S* Tone GDP				-0.00									
				(-0.97)									
P* Risk Bert					0.20								
					(1.40)								
S* Risk Bert					0.91								
					(1.31)								
P* Risk Uni						-0.40							
						(-0.48)							
S* Risk Uni						1.47							
						(0.69)							
P* Risk HHLT							-0.08						
							(-0.13)						
S* Risk HHLT							0.11						
							(0.14)						
P* Downsid Risk								1.77					
								(0.65)					
S* Downsid Risk								-3.98					
								(-0.33)					
P* Upsid Risk									13.95				
									(1.17)				
S* Upsid Risk									11.83				
									(1.60)				
P* Inflat Risk										24.62**			
										(2.63)			
S* Inflat Risk										-44.39***			
										(-5.99)			
P* Upsid Risk Inflat											145.96**		
											(2.07)		
S* Equiti Risk Premium											40.24***		
											(3.90)		
P* Downsid Risk Growth												81.39***	
												(4.55)	
S* Risk Premium Corpor												0.00	
												(.)	
Constant	0.43	2.23**	0.54	0.48	0.16	0.46	0.50	0.37	0.19	0.23	0.40	0.41	2.37***
	(0.88)	(2.40)	(1.10)	(0.99)	(0.28)	(0.73)	(0.41)	(0.68)	(0.36)	(0.47)	(0.83)	(0.83)	(3.21)
R-squared	0.44	0.52	0.49	0.47	0.46	0.45	0.44	0.45	0.47	0.48	0.46	0.45	0.53
Observations	69	69	69	69	69	69	69	69	69	69	69	69	69
P* = 0		0.67	0.02	0.10	0.17	0.63	0.89	0.52	0.25	0.01	0.04	0.00	0.24
P*= 0 and P = 0	0.19	0.37	0.03	0.07	0.20	0.33	0.42	0.27	0.14	0.02	0.08	0.00	0.09
S* = 0		0.00	0.10	0.34	0.20	0.49	0.89	0.75	0.12	0.00	0.00	.	0.00
S* = 0 and S = 0	0.58	0.01	0.23	0.53	0.39	0.67	0.89	0.82	0.18	0.00	0.00	0.54	0.00

TABLE A17—**Inflation Regressions: Romer and Romer Sample: 1979-2001**

	(1) INFL	(2) INFL	(3) INFL	(4) INFL	(5) INFL	(6) INFL	(7) INFL	(8) INFL	(9) INFL	(10) INFL	(11) INFL	(12) INFL	(13) INFL
P INFL	-0.08	0.08	-0.09	-0.07	-0.14	-0.10	0.04	-0.12	-0.15	-0.08	-0.05	-0.07	
	(-0.25)	(0.23)	(-0.25)	(-0.20)	(-0.45)	(-0.33)	(0.14)	(-0.38)	(-0.48)	(-0.24)	(-0.16)	(-0.20)	
S INFL	1.06***	0.92**	1.07***	1.03***	1.21***	1.14***	0.95***	1.12***	1.15***	1.07***	1.05***	1.05***	1.04***
	(3.29)	(2.63)	(3.02)	(3.03)	(4.07)	(3.74)	(3.04)	(3.49)	(3.55)	(3.19)	(3.30)	(3.19)	(18.01)
P* Tone LM		-0.14											
		(-0.89)											
S* Tone LM		0.02											
		(0.18)											
P* Tone Bert			-0.01										
			(-0.46)										
S* Tone Bert			0.01										
			(0.79)										
P* Tone Inflation				0.00									0.01
				(0.57)									(1.03)
S* Tone Inflation				0.00									0.00
				(1.47)									(1.04)
P* Risk Bert					0.34***								0.23
					(2.93)								(1.57)
S* Risk Bert					0.19								
					(0.66)								
P* Risk Uni						1.19							
						(1.58)							
S* Risk Uni						0.13							
						(0.09)							
P* Risk HHLT							-0.37						-0.81
							(-1.00)						(-1.67)
S* Risk HHLT							0.26						
							(0.50)						
P* Downsid Risk								1.85					
								(1.33)					
S* Downsid Risk								2.34					
								(0.65)					
P* Upsid Risk									11.40**				
									(2.19)				
S* Upsid Risk									-11.29***				-2.60
									(-2.97)				(-0.62)
P* Inflat Risk										2.69			
										(0.39)			
S* Inflat Risk										3.11			
										(0.51)			
P* Upsid Risk Inflat											100.15***		62.80
											(4.93)		(1.39)
S* Equiti Risk Premium											46.87***		60.30***
											(3.17)		(2.81)
P* Downsid Risk Growth												26.54**	
												(2.43)	
S* Risk Premium Corpor												0.00	
												(.)	
Constant	-0.16	0.28	-0.14	-0.12	-0.93***	-0.88**	0.23	-0.29	-0.25	-0.23	-0.30**	-0.17	0.47
	(-1.12)	(0.55)	(-0.73)	(-0.76)	(-3.03)	(-2.48)	(0.38)	(-1.63)	(-1.44)	(-1.11)	(-2.13)	(-1.16)	(0.52)
R-squared	0.86	0.86	0.86	0.87	0.87	0.86	0.86	0.87	0.86	0.87	0.86	0.89	
Observations	69	69	69	69	69	69	69	69	69	69	69	69	69
P* = 0		0.37	0.65	0.57	0.00	0.12	0.32	0.19	0.03	0.69	0.00	0.02	0.00
P*= 0 and P = 0	0.80	0.67	0.84	0.85	0.00	0.19	0.60	0.36	0.07	0.91	0.00	0.00	0.00
S* = 0		0.86	0.43	0.15	0.51	0.93	0.62	0.52	0.00	0.61	0.00	.	0.01
S* = 0 and S = 0	0.00	0.04	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00

TABLE A18—**Determinants of the Change in the Intended Federal Funds Rate**

	(1) OLS	(2) OLS	(3) q20	(4) q20	(5) q80	(6) q80
OLDTARG	-0.07***	-0.07***	-0.08***	-0.09***	-0.05**	-0.03
	(-4.65)	(-4.57)	(-2.90)	(-3.43)	(-2.15)	(-1.02)
P GDP	0.32***	0.30***	0.27*	0.29**	0.17	0.15
	(3.05)	(2.98)	(1.88)	(2.32)	(0.82)	(0.82)
S GDP	-0.20**	-0.20**	-0.14	-0.18*	-0.01	-0.01
	(-2.00)	(-2.10)	(-1.17)	(-1.88)	(-0.04)	(-0.05)
P Inf	0.12	0.10	0.02	0.08	0.23	0.12
	(1.10)	(0.94)	(0.13)	(0.47)	(1.10)	(0.72)
S Inf	0.03	0.03	0.14	0.08	-0.01	0.06
	(0.26)	(0.27)	(0.85)	(0.48)	(-0.03)	(0.37)
P Unem	0.09	0.05	-0.02	-0.05	-0.05	-0.05
	(0.68)	(0.34)	(-0.13)	(-0.25)	(-0.16)	(-0.16)
S Unem	-0.14	-0.11	-0.05	-0.08	-0.07	-0.04
	(-1.09)	(-0.80)	(-0.30)	(-0.40)	(-0.24)	(-0.12)
P* Risk Uni		-0.24***		-0.23*		-0.18
		(-2.67)		(-1.87)		(-0.91)
S* Risk Uni		0.24		-0.15		0.84
		(0.78)		(-0.37)		(1.62)
Constant	-0.09	0.14	-0.17	0.42	0.01	-0.11
	(-0.75)	(0.72)	(-0.51)	(1.03)	(0.06)	(-0.38)
R-squared	0.31	0.36	0.30	0.34	0.15	0.17
Observations	93	93	93	93	93	93
P* = 0		0.01	.	0.06	.	0.37
P*= 0 and P = 0	0.02	0.00	0.29	0.00	0.35	0.58
S* = 0		0.44		0.72		0.11
S* = 0 and S = 0	0.17	0.20	0.47	0.40	0.98	0.50

Notes: The dependent variable is the change in the intended federal funds rate. q20 and q80 refer to quantile regressions for the 20 percent and 80 percent quantiles. OLDTARG is the initial level of the intended funds rate. P* and S* measures are risk uni-grams. t-statistics are in parentheses; Newey and West (1987) standard errors used for OLS. Pseudo R^2 for the quantile regressions reported. P-values for the joint encompassing test that the coefficients on P^*, P, S^*, and/or S equal zero reported. The data are semiannual from 1979 to 2008. Our sample ends in 2008 to avoid running a regression with many zeros on the left-hand side.

A2. Temporal Variation and Nonlinearity in the Romer and Romer Regressions

A. Temporal Variation

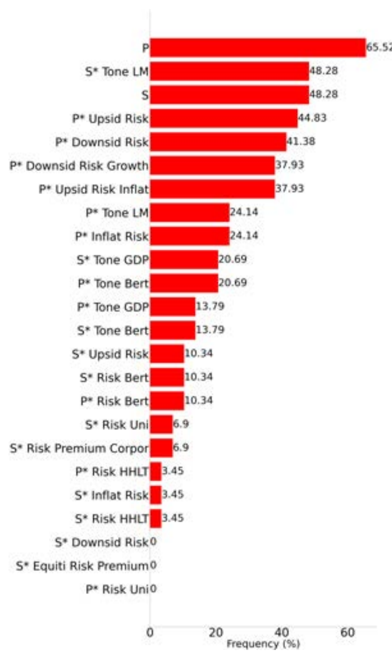
Tables 5 through 7 in the main paper identify, for OLS regressions of (7), those predictors selected by Lasso over the full sample 1979 through 2017. To investigate the degree to which the selected variables may change over time, we re-estimate the extended Romer and Romer regressions by Lasso but over rolling 10-year windows. To evidence the degree of temporal variation, Figure A3 reports the proportion of times across the 30 rolling windows that a given predictor is selected when explaining unemployment, GDP growth, or inflation. While we see from the figure that there is temporal variation, as no variable is selected 100 percent of the time, for both the unemployment rate and GDP growth we do see that as well as the point forecasts, P and S, textual measures are also commonly selected. This confirms that our main result – that the forecast narratives P* and S* add value relative to the point forecasts P and S – is generally robust to temporal instabilities. For inflation, the results in Figure A3 also bear out Table 7. While the textual indicators are selected, this happens less frequently than for unemployment or GDP growth. This is consistent with the view that the point forecasts for inflation, P and S, are closer to "pricing in" all of the information from the forecast narratives. But while there is no such dominant (over time) set of text-based measures that helps explain inflation outcomes, the text-based measures do commonly add value over and above the point forecasts.

B. Nonlinearity in the Extended Romer and Romer Regressions

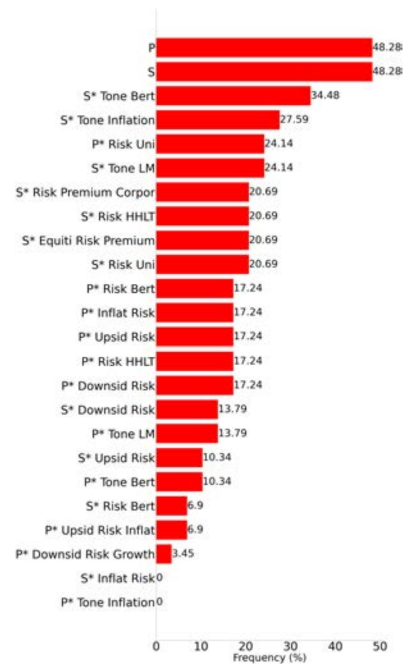
To test whether the results in the main paper are robust to P, S, P*, and S* having possibly nonlinear and/or interacted effects on the unemployment rate, GDP growth, and inflation outcomes, we re-estimate (7) using extreme gradient boosting (XGBoost). XGBoost is a machine-learning algorithm that explains the variable of interest via a sequence of decision trees and allows for nonlinearities. To summarize the importance of each variable when we re-estimate the extended regressions, (7), via XGBoost we report the Shapley (1953) values for each variable. Strumbelj and Kononenko (2010) also use Shapley-based measures to interpret fitted machine-learning models. Figure A4 shows the Shapley-based estimates of variable importance for the 3 regressions. Panel (A) shows that for unemployment P and S are the most important predictors. However, text-based measures from both P* and S* are also important, consistent with our main results in Table 5. For GDP growth, panel (B) reveals that the risk uni-gram from the FOMC is in fact even more important than either of the point forecasts for GDP growth. For inflation, in panel (C), we continue to see that, at least when explaining the conditional mean as in the OLS regression in Table 7 in the main paper, the textual measures do not add much value relative to P and S.

Figure A3. Variable Selection from Lasso

(a) Unemployment



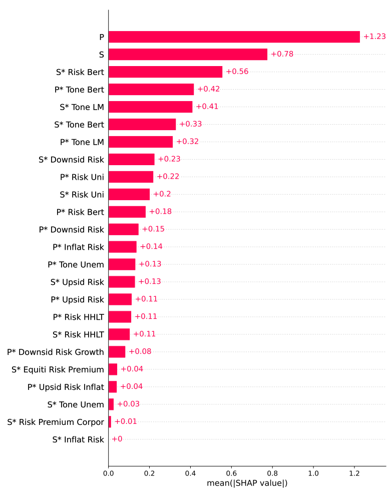
(b) Real GDP Growth



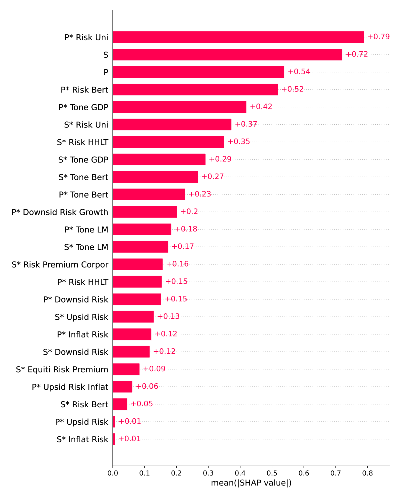
(c) Inflation

The figure displays the frequency of selected variables via Lasso. The estimation of the model is based on a 10-year rolling window. There are 30 rolling regressions estimated. Graph (A) displays results for unemployment, Graph (B) shows results for real GDP, and Graph (C) reports results for inflation. The textual measures are based on FOMC (or P*) and Greenbook (or S*) documents. The data are semiannual from 1979 to 2017.

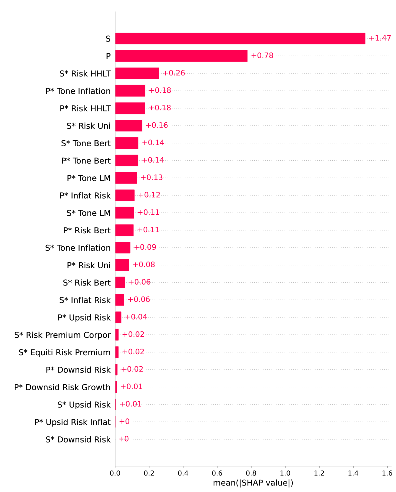
Figure A4. Variable Importance



(A) Unemployment



(B) Real GDP Growth



(C) Inflation

The figure displays the Shapley values for each predictive variable (P, S, P*, and S*) when re-estimating (7) with the extended set of textual measures by XGBoost. Graph (A) displays results for unemployment; Graph (B) shows results for real GDP; Graph (C) reports results for inflation. The data are semiannual from 1979 to 2017.