Text Shocks and Monetary Surprises: Text Analysis of FOMC Statements with Machine Learning

Amy Handlan*

September 6, 2022

Abstract

This paper shows that the wording of Federal Reserve communication affects expectations and other economic variables over and above the effects of setting the federal funds rate. Adapting neural network methods for text analysis from the computer science literature, I analyze how the wording in FOMC impacts fed funds futures (FFF) prices when these statements are announced. Using text analysis on FOMC statements and internal meeting materials, I create a new monetary policy "text shock" series for 2005-2014 that isolates the variation of FFF prices that are not generated by asymmetric information. I also find that the impact of on real interest rates is twice as large when using text shocks over other measures, like changes in FFF prices. Furthermore, the text shock produces responses in output and inflation qualitatively consistent with workhorse macroeconomic models, whereas changes in FFF prices do not.

Keywords: FOMC statements, forward guidance, machine learning, monetary policy shocks, neural network, text analysis

JEL Codes: C45, E52, E58

*Department of Economics, Brown University, 70 Waterman St., Providence, RI 02912. Email: amy_handlan@brown.edu. Website: https://sites.google.com/view/amy-handlan. I would like to express my deepest appreciation to my committee - Christopher Phelan, V.V. Chari, and Loukas Karabarbounis - for their valuable advice and support. For their helpful comments and suggestions, I also wish to thank Jacob Adenbaum, Robert Barsky, Ross Batzer, Anmol Bhandari, Jonas Fisher, Larry Jones, Ellen McGrattan, Vitoria R. de Castro, Dana Shaat, as well as seminar participants from the University of Minnesota, Brown University, the Federal Reserve Bank System, and the other institutions where I presented this paper. Finally, I thank the University of Minnesota Doctoral Dissertation Fellowship and the AEA Summer Fellows Program for their financial support.

1 Introduction

The Federal Open Market Committee (FOMC), the policy-making branch of the Federal Reserve (Fed), meets 8 times a year to discuss monetary policy and set the federal funds rate (FFR). Since May 1999, the FOMC has released a statement discussing its current and future policy objectives and assessments of US economic performance. The portion of statements that discusses future policy and future economic conditions is referred to as forward guidance. The Fed claims that they release these statements to increase transparency of monetary policy actions and provide guidance to public expectations. If forward guidance impacts the real economy by changing expectations then Fed announcements can be a viable tool to influence the economy, especially in periods when other policy tools are unavailable, like the zero-lower-bound. This raises the question: how do *the words* of FOMC statements impact expectations of future monetary policy, and more specifically future federal fund rates?

This paper studies how the information in Federal Reserve communication affects expectations and other economic variables over and above the effects of setting the federal funds rate. I adapt neural network methods from the computer science literature to use FOMC post-meeting statements to predict high-frequency changes in fed funds futures (FFF) prices. Changes in these prices encompass changes in market expectations of how the FOMC will set the federal funds rate in the future. Using the neural network, FOMC statements, and internal FOMC meeting materials, I create a monetary policy shock series called "text shocks." A positive text shock means that the FOMC announcement has shifted the path of federal-funds-rate expectations up. So, a positive text shock can also be thought of as a contractionary monetary policy shock.

This series represents innovations of monetary policy and I can use the series to study monetary policy's effect on the macroeconomy. With the new shock series, I have three main findings: first, the variation in FFF prices accounted for by the wording of FOMC statements is four times what is accounted for by changes in target federal funds rate. Second, using my text shocks instead of pure FFF price changes to represent monetary policy, I show that monetary policy has a larger effect on real interest rates compared to the literature. Third, if the FOMC releases a statement that increases the expected future value of the federal funds rate - that is, if there is a positive text shock - then output and inflation decrease, which is qualitatively consistent with a variety of macroeconomic models. Meanwhile, using changes in FFF prices as a monetary policy shocks does not produce qualitatively consistent responses.

To predict FFF price changes from FOMC statement text, I use the state-of-the-art neural network for text analysis from Yang, Dai, Yang, Carbonell, Salakhutdinov and Le (2020). A neural network is a parametric approximation of a potentially non-parametric or complex function from input to output variables (Athey and Imbens, 2019). The advantage of using a neural network is that it incorporates complex features of the text, like word order and word interdependencies,[1] for prediction tasks, like the one in this paper. Other text analysis methods, such as clustering words into topics or using word counts to create sentiment indices, have been used to study central bank communication. Creating sentiment-word lists - such as hawkish versus dovish sentiment in Lucca and Trebbi (2009), expansionary versus contractionary policy sentiment in Acosta (2022); Hansen and McMahon (2016), or degrees of policy uncertainty in Husted, Rogers and Sun (2017) - are popular because the researcher has complete control over the individual words and their interpretation. However, these word count methods often overlook more complex feature of text. Gentzkow, Kelly and Taddy (2019) acknowledge that there is room in economics for machine learning methods of text analysis. With advancements in the field for adapting neural networks to smaller datasets, these methods can be used to study the wording central bank announcements. Doh, Song and Yang (2020), the paper most similar to this one, uses machine learning to measure differences between fed announcements

[1]See Section 3 for an example.

over time and related to alternative FOMC drafts to create a measure shifts of hawkishness and dovishness over time. What makes the text shocks in my paper different is that I use the relationship between the text of the announcements and high-frequency asset prices to construct the shock.

In this paper, the neural network predicts changes in FFF prices using the joint occurrence of FOMC statement text. Rather than keep track of the entire expectations path implied by FFF prices, economists either focus on the expectations for the one FOMC meeting, as in Gertler and Karadi (2015), or use the first principal component of multiple fed fund futures price changes to capture the common variation across FFF contracts in a single dimension, as in Nakamura and Steinsson (2018) and Gurkaynak, Sack and Swanson (2004).[2] This paper uses the latter approach to represent changes in expectations of the FFR as a single output variable for the neural network.[3] Although the output variables are already numerical, the text data has to be transformed into quantitative representations before becoming inputs to the neural network. The words within FOMC statements are represented as vectors, so an FOMC statement as a whole is a matrix of numbers. These word-vectors, also called word embeddings in the text analysis literature, are such that the more similar words are the closer their vectors are. Then the parameters of the neural network are fitted, or "trained," to map the input variables to the output variable. That is, from the words of FOMC statements to the changes in FFR expectations.

Using the trained neural network, I create a new monetary policy shock series. What I call "cleaned text shocks'' are created following two steps: first, I project changes in FFF prices onto the FOMC statement text with the neural network. For

[2]Principal component analysis is a method used to reduce dimensionality of data. Using the eigenvalue decomposition of a dataset's covariance matrix, data is projected to new dimensions according to its variance. The first principal component, the first coordinate in the new dimensions, captures the greatest common variance of the original variables.

[3]While this paper uses an asset-price representation, expectations of future monetary policy are also represented with results from professional surveys, like the Survey of Economic Professionals with the Federal Reserve Bank of Philadelphia or the Blue Chip Economic Indicators and Financial Forecasts.

convenience, I will refer to the series at this stage as the "text shock" or the "uncleaned text shock." This step is to identify the changes in FFF prices that come from the announcement and not from other attitudes from trading in the futures market. Note, here I am relaxing the common assumption from the high-frequency identification (HFI) literature that assumes FFF price changes in the small window around the FOMC statement release are *only* influenced by the monetary policy announcement.

At this stage, the text shock would represent market reactions from the whole FOMC statement. However, these statements cover the FOMC's current assessment of economy in addition to monetary policy actions or guidance. If the FOMC's assessment contains more precise or superior information than what markets possess, then words other than the monetary policy action or guidance would be influencing FFF prices. This is sometimes referred to as the "Fed Information Effect." To create a series that represents new monetary policy we want to remove the information effect. Other papers, such as Romer and Romer (2004) or Bu, Rogers and Wu (2019), strip their shock series of the FOMC's private information using internal FOMC forecasts of macroeconomic variables. In a similar fashion, I look to the FOMC's internal meeting materials for the alternative versions of post-meeting statements.

The alternative statements in the FOMC meeting materials are statements that the FOMC could have released but did not. I use alternative statements for meetings from 2005-2014 that have been released to the public.[4] With the neural network trained on the actual FOMC statements, I predict the change in FFR expectations for all alternative statements for each FOMC meeting. These alternative statements all include the Fed's assessment of the current state of the economy, but they can differ in their forward guidance and policy action. To represent that common element across alternatives, I use the average of FFR expectation shifts predicted for the alternative statements for a given meeting.

The second step in producing the text shock series is then to subtract this

[4] The statements are found in Bluebooks from January 2005 - 2010 and the Tealbooks form June 2010 - December 2014. Materials are released on a five year lag.

average from expectation changes predicted from the first step. This leaves the shock series as representing monetary policy actions and not the Fed's description of the economy. I call this the "cleaned text shock." This cleaned text shock series is the change in fed funds future prices caused by the FOMC statement text and controls for the FOMC statement including non-monetary policy information that could influence markets.

This paper's contribution to the monetary policy shock literature is largely through the cleaned text shock series' representation of forward guidance. Using the wording of FOMC statements in creating the shock series is relevant for being able to study the effects of forward guidance. This is because the variation in forward guidance policy will show up as variation in wording of the monetary announcements. I test if the cleaned text shock series captures a sense of forward guidance by looking at the correlation between the shock and FFF prices at different horizons. I find that as the horizon increases, the correlation between the FFF price and the cleaned text shock increases. Also, the text shock accounts for more variation in FFF prices as the contract horizon increases. Other candidate series, such as the first principal component of FFF price changes or the uncleaned text shocks do not have these patterns across horizons.

To study the transmission of monetary policy text shocks to other economic variables, I conduct two exercises. First, I regress nominal and real interest rates at one to ten year horizons on the cleaned text shock. I compare the results with other regressions for a variety of other shock series, including the "uncleaned'' text shock, the first principal component of changes in FFF prices, the shock series from Nakamura and Steinsson (2018), and the shock series from Gertler and Karadi (2015). Nakamura and Steinsson (2018) create their shock series from the first principal component of changes in FFF prices *and* Eurodollar futures prices around the release of FOMC statements. Gertler and Karadi (2015) create their shock series as the daily change in the one-year treasury yield instrumented with high-frequency changes in the three-month-ahead FFF price around the release of FOMC statements. All of these series are measured in basis points and have similar magnitudes.

I find that all the shock series are similarly correlated with nominal interest rates, but the correlations with real interest rates are quite different. For nominal rates, I find that all of the shock series have about a one-to-one effect on nominal treasury yields. However, for real interest rates, I find the coefficients for the text shocks are about twice the size as the other shock series. For example, a one basis point increase in the cleaned text shock is associated with a four basis point increase in the two-year treasury-inflation-protected security (TIPS) yield (real interest rates). Whereas, a one basis point increase in the principal component of FFF prices, in the Nakamura and Steinsson (2018) shocks, or in the Gertler and Karadi (2015) shocks are associated with a two basis point increase in two-year TIPS yields. From this exercise, I conclude that monetary policy has a larger effect on real rates than other HFI monetary shocks would indicate because the cleaned text shock is able to use the wording of FOMC announcements to capture the effect of forward guidance.

For the second exercise I use an external instrument, vector autoregression (VAR) approach to study the relationship between monetary policy shocks and other macroeconomic variables. As in Gertler and Karadi (2015), I include industrial production, Consumer Price Index (CPI), one-year treasury yield, and an excess bond premium measure in the estimation. I use the local projection method from Jordà (2005) to produce impulse response functions. An increase in the cleaned text shock series is associated with responses in output, inflation, and excess bond premium that are consistent with workhorse macroeconomic models. That is, when using the text shock to represent monetary policy innovations, a contractionary shock produces decreases in output and inflation and an increase in the excess bond premium. This indicates that forward guidance through the Fed's wording of FOMC statements is an important channel to qualitatively match how monetary policy impacts the economy in data and in a variety of macroeconomic models.

I compare these impulse responses to the responses of macroeconomic variables to other monetary shock series, such as the change in the three-month-ahead FFF price from Gertler and Karadi (2015) and the principal component of changes in multiple FFF prices. As in Ramey (2016), I find that in the local projection framework that using FFF prices alone as the monetary shock produces minimal responses in output, inflation, and excess bond premium variables. In fact, a contractionary Gertler and Karadi (2015) shock is associated with slight increases to inflation and output when using the local projection estimation.

Overall, this paper seeks to address the following question: how do monetary policy announcements affect the economy by influencing expectations of future monetary policy action? There is a long literature focusing on this question. When looking at financial markets, Ai and Bansal (2018), Cieslak and Schrimpf (2019), and Lucca and Moench (2015) all find that FOMC announcements, measured with changes in FFF prices, have sizable influences on bond risk premia. This means they find that the Fed influences investor expectations of the future path of the economy. Through this channel, Campbell, Evans, Fisher and Justiniano (2012), Gertler and Karadi (2015), Gurkaynak, Sack and Wright (2007), Kuttner (2001), Lucca and Trebbi (2009), and Nakamura and Steinsson (2018) show the effect of monetary policy announcements using structural models or VAR methods. Although these papers agree that announcements can have sizable impacts, there is often disagreement on the qualitative direction of announcement effects. For example, Campbell et al. (2012) find a self-fulfilling, or "delphic", effect of forward guidance where increased expectations of the FFR are associated with increased output growth and decreased inflation. Conversely, using external instruments, SVAR approach, Gertler and Karadi (2015) find increased FFR expectations lead to the opposite. Without FFF contracts, other papers, like Romer and Romer (2004) and Bu et al. (2019), argue that accounting for the "Fed Information Effect" is important for producing monetary shock series that have "correct" impulse responses. My paper contributes to this discussion with my text shock series, which is derived from variation in FOMC statement text and accounts for the "Fed Information Effect." The main result of this paper is that monetary policy does have an impact on the economy and unconventional tools, like forward guidance, are important for qualitatively matching theoretical effects of monetary policy from workhorse macroeconomic models.

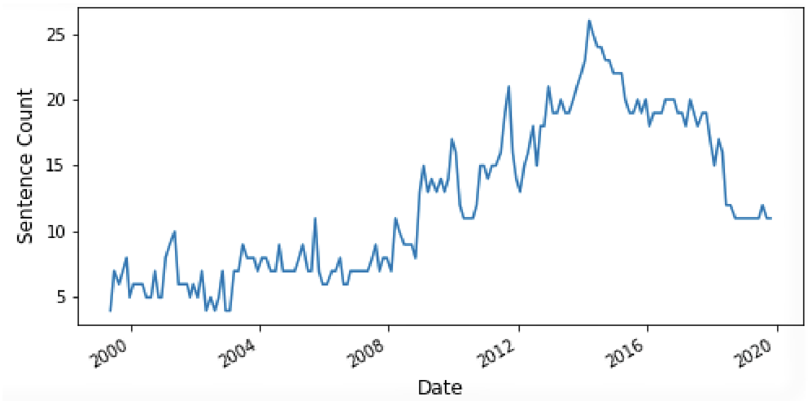
The rest of the paper proceeds as follows: Section 2 details data sources and preparation methods. Section 3 describes the text-analysis method and results for predicting monetary surprises with FOMC statement text. Section 4 describes the creation of the new monetary policy shock series, text shocks. Next, section 5 includes comparisons of my shock series with others from the literature. And in section 6, I conclude.

2 Data

The sample period for my analysis is from May 1999 through October 2019. The FOMC post-meeting statements during this period were sourced from the Board of Governors of the Federal Reserve System website. Handlan (2020) contains a table with all 165 statements with their date and time of release. I drop unscheduled FOMC meetings' statements from the sample. This is because I want the change in asset prices that occurs around the statement's publication to be from the content of the statement, not a combination of the statement wording and the surprise that there was meeting.

These statements generally discuss the current economic environment, the new target federal funds and discount rate, and information about the FOMC's expectations for the future of the economy. Following the 2008 Financial Crisis, the statements also discussed unconventional monetary policy, such as quantitative easing programs. This added topic and the inability of the FOMC to use changes in rates to influence expectations increased the length of statements post-2008, as seen

Figure 1: Number of Sentences in FOMC Statements, 1999-2019



Note: The above counts are for FOMC statements that have already been cleaned, described in the appendix

in Figure 1.

When creating the cleaned text shock series, I will use alternative versions of FOMC statements that could have been released but were not. Alternative statements are provided to FOMC members in their pre-meeting materials. Pre-meeting materials that are sent to FOMC members before the policy meetings describe the state of the economy and recommend policy actions. These materials are bundled into books. Since 2010, that has been the Tealbook A and B. Previously there were Greenbooks and a Bluebook. These books are released to the public on a five year lag and are also available on the Board of Governors of the Federal Reserve System website. Drafts for alternative versions of FOMC post-meeting statements are clearly displayed in the Tealbook B's and in the Bluebooks from January 2005 through December 2014. Prior to 2005, wording for statement drafts is spread-out throughout the book and is not clearly labeled. Accordingly, I limit the sample to statement alternatives that are written in their own section and clearly labeled in the pre-meeting materials.

Tick-level time-of-sale data on federal fund futures at the one to six month horizons was purchased from CME Group. I have this price data for the entire sample of May 1999 through October 2019. As is common in monetary economics, I use FFF prices as a proxy for market expectations of how the FOMC will set federal funds rate at future meetings. This stems from how the FFF contract is priced: the contract settlement price is determined by the average effective federal funds rate over the final month of the contract. From prices, I calculate the change in market expectations of the FFR for the current and next FOMC meeting. The changes in expectations of the federal funds rate at the current meeting and for the next FOMC meeting are highly correlated.[5] To only focus on a single dimensional representation of expectations, I use the first principal component of these two variables as the baseline expectations representation.

Following Nakamura and Steinsson (2018), I re-scale this principal component such that a one-unit increase corresponds to a 100 basis point increase in the daily change of a one year treasury yield. Translating the first principal component back to changes in federal funds rate expectations, a one unit increase in the first principal component would translate to the federal funds rate being set 168 basis points higher than expected and a 180 basis point increase in the expectations for the federal funds rate at the next FOMC meeting. For the rest of the paper, I will interchangeably refer to this measure as the change in federal funds rate expectations or the change in fed funds futures prices.

Throughout the paper I also use data on the target federal funds rate, treasury yields, industrial production, consumer price index (CPI), and excess bond premium. Daily target federal funds rate data are pulled from FRED. When the target federal funds rate is a range of values, I take the average of the of range to get a single number representation of the target federal funds rate. Monthly data on industrial production and CPI are also collected from FRED. Daily data for treasuries and Treasury Inflation Protected Securities (TIPS), representing nominal and real interest rates respectively, are from Gurkaynak et al. (2007) and Gürkaynak, Sack and Wright

[5] Figure C1 graphs this correlation.

(2010), receptively. Both data sets are available on the Federal Reserve Board's website.[6] I use the monthly measure for excess bond premium from Gilchrist and Zakrajsek (2012), which is also available on the Fed's website.[7]

3 Text Analysis of FOMC Statements

I use an off-the-shelf neural network for text analysis from the computer science literature to approximate the mapping from FOMC statement text to changes in federal funds rate (FFR) expectations. I use the pre-trained neural network from Yang et al. (2020), called XLNet. In this section, I will give an overview of the text analysis neural network from Yang et al. (2020). I then will discuss the application of their method to FOMC statements. Next will be an evaluation of the fitted neural network's ability to predict unanticipated shifts in FFR expectations compared to using changes in the target rate to predict expectation shifts. The section will wrap up with examples how changes in FOMC statement wording changes the neural network predictions to shed some light the nuances the neural network picks up.

3.1 Neural Network for Text Analysis (XLNet)

Innovations in computer science and text analysis, specifically "transfer learning," are allowing machine learning algorithms to be applied to smaller data sets. So now that it is possible to apply these more advanced methods to central bank announcements, why would we? The main advantages of these neural networks for text analysis is that they can capture a sense of context from words using word order and relationships between words throughout sentences.

Counting words from particular sentiment-lists or word clustering often miss how words within a sentence relate to each other. For example, the phrases "*inflation*

[6]Treasuries: https://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html, and TIPS: https://www.federalreserve.gov/pubs/feds/2008/200805/200805abs.html

[7]https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/files/ebp_csv.csv

went up, but *employment* did not" and "*employment* went up, but *inflation* did not" would produce the same measures. Methods that use neighboring words, called n-grams, would also miss information content when concepts are spread out throughout a sentence. For example, a bigram looks at the frequency of sequential word pairs. For the following sentence, "economic growth slowed, but is likely to expand at a rapid pace," a bigram would count "growth slowed'' but would miss that the point of the sentence is that economic growth is expected to increase.

Because the neural networks are able to account for nonparametric relationships between words, it has a way to approximate context and pick up long-term dependencies. Consider again the sentence, "economic growth slowed, but is likely to expand at a rapid pace." With neural network methods, "expand at a rapid pace'' and "slowed'' could both be associated with "economic growth'' for prediction even though those words are not adjacent in the sentence. Gentzkow et al. (2019) acknowledge that there is room in economics for machine learning methods of text analysis. With advancements in the field for adapting neural networks to smaller datasets, these methods can be used to study central bank announcements. This paper is the first to so.

The neural network from Yang et al. (2020), called XLNet, is considered a state-of-the-art method for text analysis tasks like translation, question and answer, classification, and regression. In other words, their method is flexible enough to approximate mappings from text to text, text to categories, or text to continuous numbers. At a high level, this algorithm is able to translate words into numerical vectors that can be aggregated to represent word use in text documents. Given this numerical "understanding" of text, we can adapt the model with a little additional finetuning for downstream tasks like prediction. For more detailed information on the text analysis, neural networks, and the XLNet model specifically are in the online appendix.

3.2 Application to Monetary Statements and Expectations

In this paper, I use the network structure and pre-trained word representations from Yang et al. (2020). Furthermore, I use the pre-trained base XLNet parameters as initial values for my task: predicting changes in FFR expectations from FOMC statement text. To prepare the text for training, I remove URLs, time of release, and voting records of FOMC members from the statements. A more detailed description of text cleaning and algorithm for training are available in the online appendix.

I split my sample into training and testing samples such that 20 percent of the sample is in the testing set. I condition splitting the meeting observations on how the target federal funds rate changed, who the Fed Chair was, and if the date was pre- or post-2007. As is common in machine learning, I train the neural network for different training/testing splits. One way to think of this is by splitting the data into five subsets. Then the network would be trained five different times where each training would correspond to one of the five subsets being assigned as the testing sample and the remaining observations would be the training sample. This is called a "k-fold cross validation," which is similar to "leave-one-out cross-validation (LOOCV)'' but with more than one observation left out of training. The results for this version of the paper are for one such training-testing split.

I fine-tune the pre-trained neural network from (Yang et al., 2020) to predict changes in FFF from the text of a FOMC statement. As is the standard in machine learning spheres, the metric for evaluating the performance of the trained network is based on how well it can accurately predict changes in FFF for FOMC statements that were not used to train the network parameters. In their summary of machine learning in economics, Athey and Imbens (2019) comment that evaluation of neural network models is inherently different from traditional econometric models. For the former the emphasis is on ability to predict outcome variables given input variables. Meanwhile, the latter focuses on estimating parameters that are functions of the joint distribution of data, construct confidence intervals of those estimates, and rely on theoretical foundations for efficiency of those estimators. Because I use a neural network approach, I will be using that literature's method for evaluating the results.

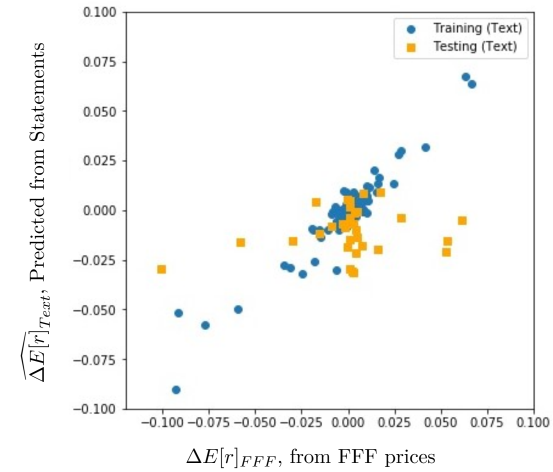
Individual parameters of the neural network are not interpretable in the same way as estimators from parametric models (Athey and Imbens, 2019). Accordingly, it is not possible to interpret the effect of one word or phrase over another. The primary goal of this paper and these text-analysis methods is to first see if we can make accurate predictions by approximating complex functions. Although I will not be able to explain the causal mechanism behind the FOMC-statements-market-expectations relationship with the trained neural network, I will be able to approximate the relationship and use prediction to create quantitative measures to describe Fed communication over time.

3.3 Evaluation of Neural Network Prediction

Neural networks are different from traditional econometric models because their evaluation is based on their ability to predict out-of-sample data, that is, data that was not used to train the neural network weights (Athey and Imbens, 2019). To evaluate the prediction, I use the Pearson correlation between the predicted output and the actual output values for the testing data. Figure 2 graphs the actual $\Delta E[r]$ on the horizontal axis against the $\widehat{\Delta E[r]}$ predicted from FOMC statement text through the neural network. The blue circle dots are the training sample while the orange squares are the testing sample. The testing sample's prediction has a 20% correlation with actual $\Delta E[r]$ data. The training sample has a much higher prediction accuracy because the neural network weights change to match $\widehat{\Delta E[r]}$ and $\Delta E[r]$. Together, plotting the training and testing data, I there is a 72% correlation between the actual data and the neural network output. Figure C2 graphs $\widehat{\Delta E[r]}$ and $\Delta E[r]$ over time.

The large difference between in-sample and out-of-sample accuracy could mean that there is overfitting of the neural network parameters to the training data.In machine learning, there are a few procedures to minimize overfitting problems and to

Figure 2: Neural Network Prediction on Training and Testing Samples



Note: $\Delta E[r]_{FFF}$ is the first principal component of two variables: changes in expectations of the federal funds rate for the current meeting and the next meeting. These expectations are calculated from changes in FFF prices from 10 minutes before to 20 minutes after the FOMC announcement is released. The vertical axis,$\widehat{\Delta E[r]}_{Text}$, is the neural network's prediction of $\Delta E[r]_{FFF}$ from FOMC statements. The scale is such that 0.025 on the horizontal axis represents a change in $\Delta E[r]$ by 2.5 basis points.

make the neural network a more accurate approximation of the mapping from inputs to outputs. In the computer science literature this would phrased as trying to make the neural network more more generalizable to new data.

What I did first was limit the network training through the learning rate and number of training iterations. When there are too many training iterations, the researcher can see that eventually the out-of-sample prediction accuracy stops increasing and begins to decrease. This is a sign of the network weights overfitting the training data. When the learning rate is too high, this degradation of out-of-sample prediction happens more quickly, meaning over fewer training iterations. Because I am using a transfer learning approach, the parameters of the neural network start out as having a general weighting scheme for interpreting words in text. So limiting how much the parameters can update, either within each iteration or over all iterations, would help keep the weighting more generalizable. I also tracked the out-of-sample accuracy while training occurs and stopped the training of the network once out of sample prediction decreases from the previous iteration. This intuitively lead to an decrease for the in-sample prediction accuracy. The importance is to strike a balance between teaching the network about the desired mapping - from FOMC statements to changes in expectations - and training the network to the point of overfitting.

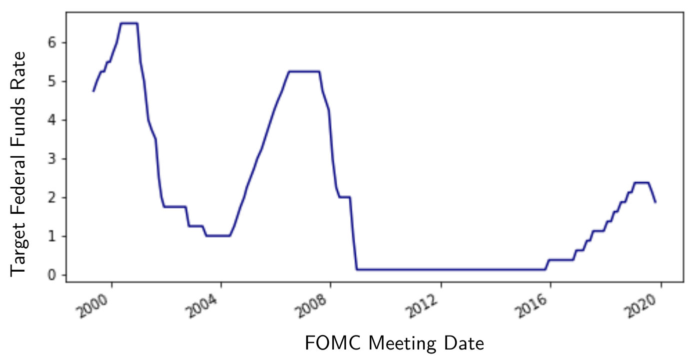
The second robustness check is to find more training data. In terms of machine learning problems, 165 observations is incredibly small even with a transfer learning approach. However, there are only so many FOMC statements that have been made over time. One approach is to artificially augment the training sample with a method called 'back-translation.' Computer scientists have shown that translating text inputs to a different language and then translating them back to the original language with a software like Google Translate can create synthetic training observations that improve network performance. The underlying assumption is that Google Translate will create small variations in word order or word choice without dramatically changing the tone or content of the text. Accordingly, the back-translated statement can be assigned the same change in expectations, the same output variable, as the original FOMC statement it was created from. Preliminary results from this robustness check produce similar results as above. Accordingly, I proceed with the neural network trained on the 132 FOMC statements.

Therefore, I interpret these differences as representing changes in expectations that were not caused by the announcements. It is possible that this is due to the networks' poor ability to approximate the underlying mapping from FOMC statement to expectation changes. Because there is no standard theory in the computer science literature to verify the quality of a neural network other than out-of-sample prediction and cross validation, I complete the following exercise to put the accuracy of the neural network in a more familiar context.

3.4 Changes in Expectations and the Target Policy Rate

In monetary policy announcements, the FOMC announces their target for the federal funds rate (FFR). Following the 2008 Financial crisis when the Fed set the FFR to zero for an extended period of time and yet federal funds futures (FFF) prices still fluctuated. This highlighted that monetary policy extends beyond setting the target rate. Figure 3 shows how the target FFR has changed over the sample period. To put the neural network's prediction in context, I ask: how well do announced changes in the target FFR do in predicting shifts in expectations?

Figure 3: Target Federal Funds Rate Over Time, 1999-2019



Note: Daily target federal funds rate data is pulled from FRED. When the target federal funds rate is a range of values, I take the average of the of range to get a single number representation of the target federal funds rate.

I compare the predictive power of changes in the target rate on FFR expectations with the neural network predictions from FOMC statement text. For the former, I regress changes in FFR expectations on the change in the target rate.

$$\Delta E_t[r] = \beta_0 + \beta_1 \Delta TargetFFR \tag{1}$$

To compare apples to apples, I estimate the Equation 1 on the same observations I used to fit the neural network parameters. Increasing the federal funds rate by 0.25, the common increment for changes in the target rate, is associated with a 0.015, about one half standard deviation, increase in first principal component measure of FFR expectations.

Using this coefficient, I calculate the predicted change in expectations for observations in the testing data. I use out-of-sample prediction accuracy to compare the prediction power of the neural network with FOMC statement text versus the regression with change in the target federal funds rate. I report both the Pearson correlation between the predicted change in expectations $\widehat{\Delta E[r]}_{TFFR}$ compared to the change of expectations $\Delta E[r]_{FFF}$ implied by FFF prices with the corresponding R^2 value in Table 1. I find that the statement text can better predict changes in expectations out-of-sample than using the target rate. When looking at correlation, the statement text is twice as accurate. When considering the R^2 measure, the statement text can explain four times the out-of-sample variation compared to predictions from changes in the target FFR.

Table 1: FOMC Statement vs. Target Rate Out-of-Sample Prediction

	FOMC Statement Text	Δ Target FFR
Correlation ($\widehat{\Delta E[r]}$,$\Delta E[r]$)	0.2	0.1
R^2	0.04	0.01
N	33	33

Note: Parameters for each prediction method are fitted to the training data. Then those parameters are used to predict changes in expectations for the testing sample (N=33). The correlation in the top row is between the change in FFR expectations calculated from FFF, $\Delta E[r]$, and the predicted changes in expectations,$\widehat{\Delta E[r]}$, from either the FOMC statement text with the neural network or the changes in the target FFR with an OLS regression.

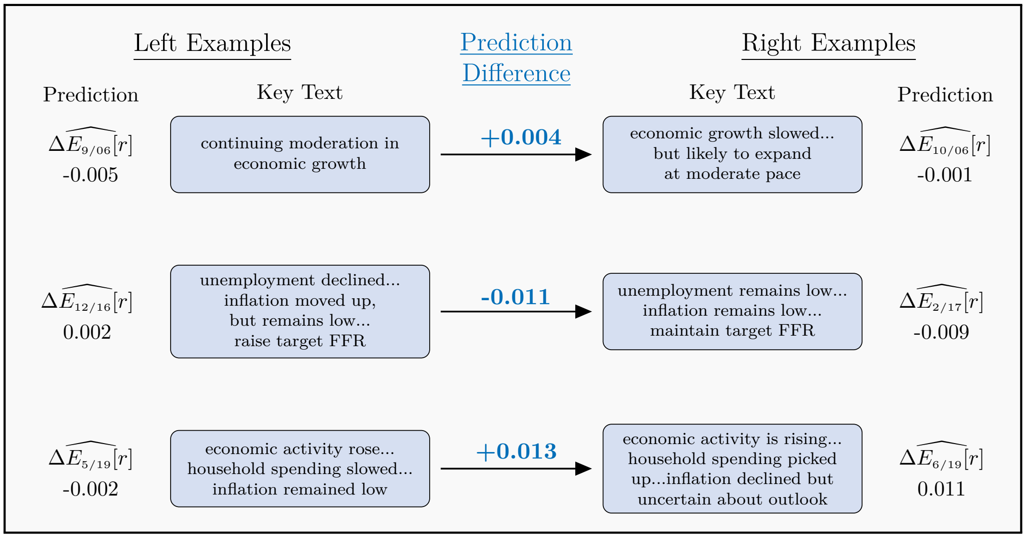
This difference is likely due to the FOMC statement contains the multiple dimensions of information. Accordingly, unexpected changes in expectations are responding to this new information and the statement text is able to capture more than just the single dimension of the federal funds rate target. Ultimately, this comparison is meant to convey, that even though the neural network is not able to strongly predict changes in expectations out of sample, it can do better than traditional measures of monetary policy changes that are publicly available when FOMC statements are released.

3.5 Different Wording Leads to Different Predictions

To shed some light on the neural network predictions, in this section I include examples of what the neural network predicts for different FOMC statements. The main exercise is to show two statements that are identical except for small differences in the text. Then to look at difference in what the neural network predicts for the statements' corresponding FFR expectation shifts. In Figure 4, there are three statement pairs for comparison. The text in the graphic highlights the text that is different between the two statements in each row. Longer excerpts of the text are included in the online appendix.

The first row compares the September 2006 statement with the October 2006 statement. Both imply that economic growth is currently slow. Both kept the target federal funds rate unchanged. But the October 2006 statement adds that the FOMC expects the economy to expand. As a naive reader, knowing nothing else besides this difference, one would expect the Octboer 2006 statement to increase federal funds rate expectations more than the September statement because traditional monetary theory indicates that inflation follows economic growth, which would trigger contractionary policy action and the Fed would increase the FFR. Other text analysis methods, such as bigrams or trigrams which look at occurrences of neighboring two or three words, would likely identify these two statements as identical. One of the strengths of the neural network method is that it can pick up on relationships between words that are connected even if the words are not literally next to each other. This shows up as a difference in predictions of 0.004. The number is about one half of a standard

Figure 4: Neural Network Prediction for Different FOMC Statements



Note: Each row is comparing two FOMC statements that have very similar text. The differences are what is written in the blue boxes. These are example statements have few differences in wording as to be able to identify what is causing the change in neural network predictions.

deviation.

The second row compares the December 2016 and February 2017 statements. Moving from the December 2016 to February 2017 statements, the main differences are going from a notion of higher inflation to one of lower inflation. We would expect that statements that discuss low inflation and no target federal funds rate changes to have a relatively more negative change in expectations compared to a statement that raises the target rate and discusses increasing inflation. The network also picks up this difference.

The final row compares the May 2019 and June 2019 statements. This comparison shows how the FOMC's confidence in their guidance impacts the neural network prediction. Both the May 2019 and June 2019 statements talk about increases to economic growth and low inflation. However, the June 2019 statement qualifies its prediction of inflation going forward. Moving from a statement where inflation is likely to stay low to a statement where the FOMC is uncertain about the path of inflation would likely encourage market to expect possible increases in federal funds rates in the future to match the potentially rising inflation. The neural network shows this as a positive difference in switching from the May 2019 to the June 2019 statements.

4 Monetary Policy Text Shocks

Federal funds rate (FFR) expectations are often measured with fed funds futures (FFF) where the change in FFF prices represent unanticipated change in monetary policy. This comes from the the efficient market hypothesis, that says all publicly available information is incorporated into asset prices. So changes in the asset prices in a small time window represent incorporation of new information into prices. In terms of fed funds futures, whose pricing structure is based on FFR expectations, a change in prices represents unanticipated changes in FFR expectations. If the change was expected, then the futures price would not have changed.

In papers like Gertler and Karadi (2015), FFF price changes themselves are used as a proxy for structural monetary policy shocks. Timing restrictions for evaluating FFF prices mean that change in fed funds futures prices in a small window around when the FOMC announcement release are mostly caused by FOMC announcement. However, other factors – such as market momentum or attitudes of traders – can impact asset prices even in that small window (Lucca and Moench, 2015; Neuhierl and Weber, 2018). And the researcher must separate the effect of exogenous shocks and the Fed's policy response to the state of the economy. Gertler and Karadi (2015) regress changes in FFF on internal economic forecasts in the FOMC's meeting materials and use the residual from this regression as the exogenous shock. This cleaned-up shock has minimal impact on economic variables.

When creating my new monetary policy shock measure, I use the FOMC statement text and the trained neural network to isolate the changes in FFF that are coming from announced monetary policy. A projection of the change in fed funds futures prices directly onto the wording of the FOMC statement looks at the change in federal funds rate expectations (measured with fed funds futures prices) explained by the monetary policy announcement itself. This projection separates other market effects on expectations from the effect of monetary policy shocks.

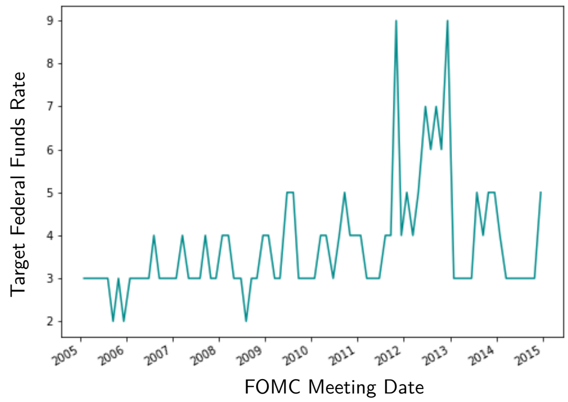
$$\textbf{Text Shock}_t = \widehat{\Delta E_t[r]}_{released} \tag{2}$$

However, this measure still has the issue that it is capturing both revelation of more precise information about the current economic situation and the monetary policy action. To have a measure of monetary policy, the former component must be separated out (Gertler and Karadi, 2015; Ramey, 2016; Romer and Romer, 2004).

To do this, I use predicted changes in FFR expectation for alternative FOMC statements to control for potential statements the FOMC could have released. Alternatively worded statements are included in the FOMC's meeting materials, called the Tealbooks and Bluebooks. These materials include information about about economy, forecasts, and policy recommendations. These books are sent to FOMC members at least one week before FOMC meeting takes place. However, the books are only released to the public on a five year lag. Figure 5 shows the number of alternative statements from 2005-2014, the period in which I have access to clearly identifiable alternative statements.

The actual statement and the alternative statements in FOMC meeting materials were all drafted with the same information. I represent the Fed's private information as the average of predicted expectation changes from alternative statements. So, the difference between the average change in expectations and the change in expectations from the actual statement that was released is the cleaned proxy for structural monetary policy shocks. This "cleaned monetary policy text shock" is the series I use in later analysis. For each FOMC meeting in this period, I calculate the

Figure 5: Number of Alternative Statements per FOMC Meeting



Note: Alternative statements are collected from the Tealbooks and Bluebooks for the FOMC meetings from 2005 to 2014.

monetary policy text shock according to Equation 3:

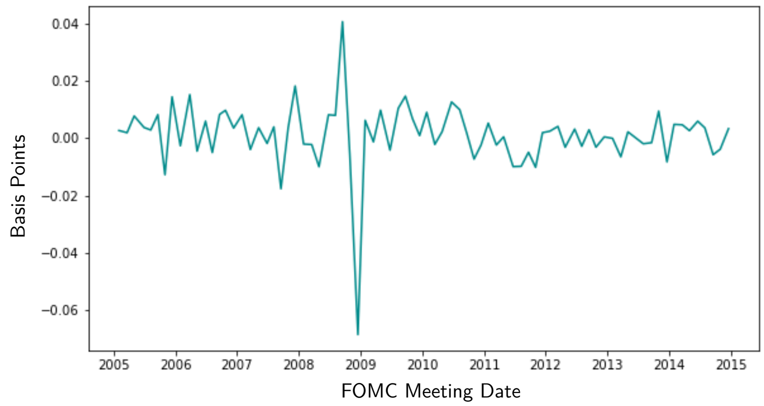
$$\textbf{Cleaned Text Shock}_t = \widehat{\mathbb{E}_t[r]}_{released} - \frac{1}{|Alts_t|} \sum_{i \in Alts_t} \widehat{\mathbb{E}_t[r]}_i \tag{3}$$

where t indexes FOMC meetings, i indexes the alternatives statements among the collection of alternatives at meeting t: $Alts_t$. The statement that was actually released is indexed as $i = I$. Therefore, $\widehat{\mathbb{E}_t[r]}_I$ represents the projection of the change in FFR expectations onto the actual FOMC statement and $\widehat{\mathbb{E}_t[r]}_i$ is the counterfactual change in FFR expectations for alternative i. I create this shock series for every meeting in January 2005 through December 2014. This date range is limited by the availability of FOMC meeting materials that contain the alternative statements.

5 Comparison with Other Monetary Shock Series

In this section, I will compare the text shock and cleaned text shock series to other monetary policy shocks from the literature. I summarize the names, notation, and description of each of the monetary shock series I will be working with in Table 2. All of the following shock series are based, at least in part, on high-frequency identi-

Figure 6: Cleaned Text Shock



Note: This shows the cleaned text shock series from 2005-2014. This is calucated in two steps, first: the first principal component of changes in expectations of the federal funds rate for the current meeting and the next meeting are predicted by the FOMC statement text with the neural network. Second, for each meeting I subtract out the average of predicted changes in FFR expectations from every alternative statement. This produces a shock series that focuses on changes in fed fund futures predicted from the FOMC statement wording and it controls for the FOMC's private information.

fication arguments. Table B1 includes summary statistics of each of the shock series. The main takeaway from that table is that the series' ranges are all very similar. Accordingly, differences in coefficient magnitudes in the subsequent sections is begin driven by what these shocks represent, not scaling differences.

In the following subsections, I will compare the effects of the new text shocks with other shock series from the literature. First, I will show what theses series say about monetary policy's effect on nominal and real interest rates. Then I will estimate different impulse response functions to show what these shock series reveal about monetary policy's effect on other macroeconomic variables.

5.1 Nominal and Real Interest Rates

In this section, I compare the effect of monetary shock series on nominal and real interest rates at different horizons. The daily change in Treasury yields represent the change in nominal interest rates. The daily change in TIPS yields represent the change in real interest rates. The daily change is calculated on the end-of-day yields for the

Table 2: Monetary Policy Shock Series

Series Name	Abbreviation	Description
Cleaned Text Shocks	$\widehat{\Delta E[r]}_{clean}$	Predicted effect on expectations from FOMC statement wording stripped of Fed's private information
Text Shocks	$\widehat{\Delta E[r]}_{text}$	Prediction from FOMC statement input into the trained neural network
PC1 FFF Price Changes	$\Delta E[r]_{FFF}$	First principal component of change in fed funds futures ($\Delta E_t[r_t]$, $\Delta E_t[r_{t+1}]$)
Gertler and Karadi (2015) Shocks	GS Shock	Daily change in 1-year treasury yield instrumented with change in 3-month fed funds future (FF4)
Nakamura and Steinsson (2018) Shocks	NS Shock	First principal component of change in fed funds futures ($\Delta E_t[r_t]$, $\Delta E_t[r_{t+1}]$) and Eurodollar futures at 2,3,4 quarters

Note: Summary statistics for the monetary policy shock series are in Table B1

day before to the day of the FOMC announcement. Table B2 and Table B3 include the summary statistics for the interest rate changes. The change in fed funds futures used to calculate the shock series occurs within a smaller, nested event window of the treasury and TIPS yield changes. This timing restriction implies the daily change in treasuries is not impacting the regressors.

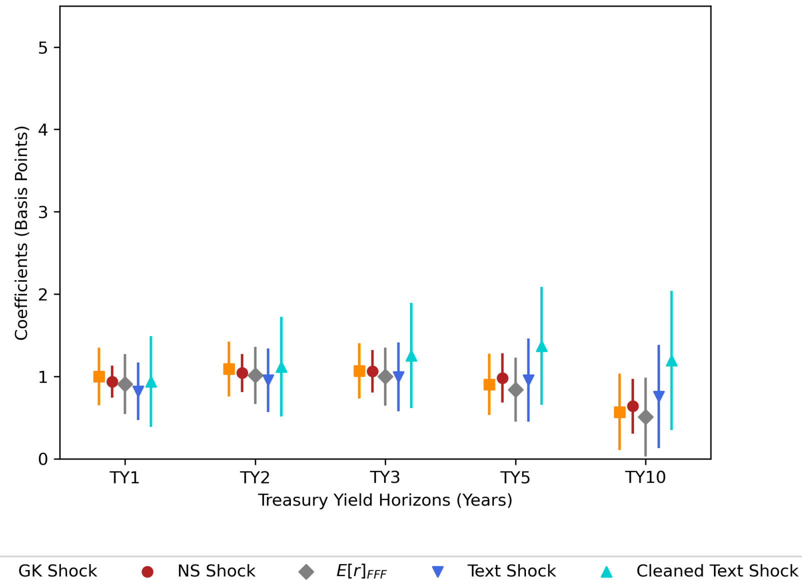
The regression specification is as follows:

$$\Delta Yield^{\ell,i} = \beta_0^{\ell,i,k} + \beta_1^{\ell,i,k}\ (monetary\ shock)^k + \varepsilon^{\ell,i,k} \tag{4}$$

where ℓ indicates either Treasury or TIPS yields, i is the horizon of the yield, ranging from 1 year to 10 years, k indexes the shock series from Table 2. The regression results are summarized in coefficient plots for nominal interest rates in Figure 7 and of real interest rates in Figure 8. Regression results for each (ℓ, i, k) specification shown these plots are available in table form in the online appendix.

A large cleaned text shock means that the predicted effect of the released

Figure 7: Nominal Interest Rates and Monetary Shocks

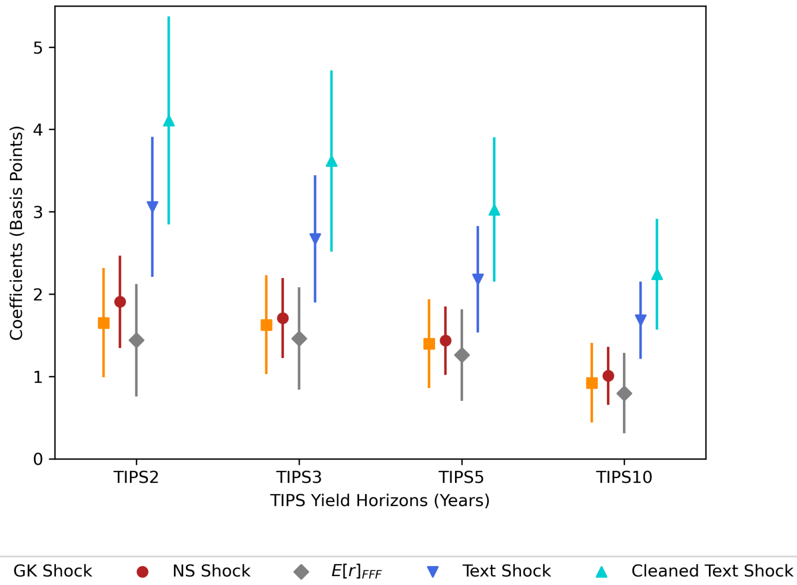


Note: The dots represent coefficients for different OLS regressions. Standard errors are Newey and West (1987) standard errors. The time sample is for 2005-2014 for all regressions.

FOMC statement is substantially different from the predicted effect of all alternative statements. The text shock and the cleaned text shock have similar effects on nominal interest rates compared to the NS Shocks and GK Shocks.

However, the text shocks have a much larger correlation with real interest rates compared to GK Shocks or NS Shocks. The coefficient is approximately double. The summary statistics show that the range of these shock series are similar, so the differences in coefficients is not driven by variability in scales across the shock series. I argue that the projection of asset prices onto the FOMC statement text is the important difference. To interpret this graph would be that the text shocks are picking up a larger effect of monetary policy on the real economy through the expectations channel compared to the literature.

Figure 8: Real Interest Rates and Monetary Shocks



Note: The dots represent coefficients for different OLS regressions. Standard errors are Newey and West (1987) standard errors. The time sample is for 2005-2014 for all regressions.

5.2 Impulse Responses with Local Projections

To study the transmission of monetary policy announcements to other variables in the economy I estimate impulse responses with a local projection and external instrument approach. As in Gertler and Karadi (2015), I include log industrial production, log consumer price index (CPI), one-year treasury yield, and excess bond premium. The excess bond premium is from Gilchrist and Zakrajsek (2012) and represents the risk premium from the difference between private and public bonds. Incorporating this variable in the specification allows the monetary shock to influence economic variables through financial markets. Summary statistics for these variables are in the appendix in Table B4.

$$Y_t = [g_t, \pi_t, ty_t, ep_t] \tag{5}$$

and g_t is the natural logarithm of industrial production, π_t is the natural logarithm of the Consumer Price Index, ty_t is the the 1-year treasury yield, and ep_t is the excess bond premium from Gilchrist and Zakrajsek (2012). I index the elements of Y_t by i.

Structural monetary shocks are then represented by $\epsilon_{2,t}$, where $Y_{2,t} = ty_t$. Because this is not measurable, economists use a proxy Z_t for these shocks such that:

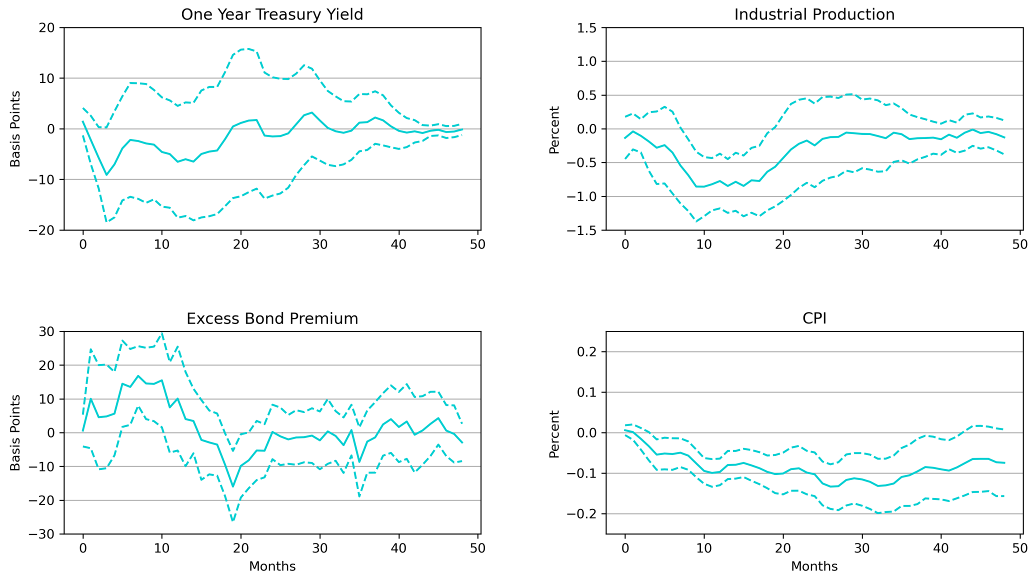
$$\mathbb{E}[z_t \; \epsilon_{2,t}] \neq 0 \qquad \mathbb{E}[z_t \; \epsilon_{-2,t}] = 0 \tag{6}$$

As discussed in an earlier section, I argue that the cleaned monetary policy text shocks meet this condition while other shock series that are created only from FFF prices are likely violating this condition. Nevertheless, to contextualize the responses of Y_t variables to an increase in the cleaned text shock, I also estimate responses to an increase three-month-ahead FFF contract as an extension of the shock series from Gertler and Karadi (2015). The different shock series are indexed by k.

For all variables to be measured on the same frequency, I convert shocks to a monthly frequency such that months without FOMC meetings have a monetary policy shock of zero. Table B5 includes summary statistics for the monthly shock series. In Gertler and Karadi (2015), the shock series is converted to a monthly series by using a rolling average so that even months without FOMC meetings can have non-zero monetary shocks. However, I use changes in the 3 month ahead FFF (FF4) to calculate the GK shock series without the rolling aggregation for comparison with my cleaned text shocks. This change does not change the results much quantitatively and the qualitative takeaways remain.

I use the local projection method from Jordà (2005) to graph impulse response functions for the different shock series and for the components of Y_t. Therefore, I run a separate regression for each shock, indexed by k, and component of Y, indexed by

Figure 9: Impulse Responses to Cleaned Text Shock



Note: Impulse responses are calculated using the local projection method from Jordà (2005). Confidence bands are at the 90% level. Standard errors are Newey and West (1987) standard errors that correct for heteroskedasticity and autocorrelation. The above are responses to a 100 basis point increase in the shock series (not pictured).

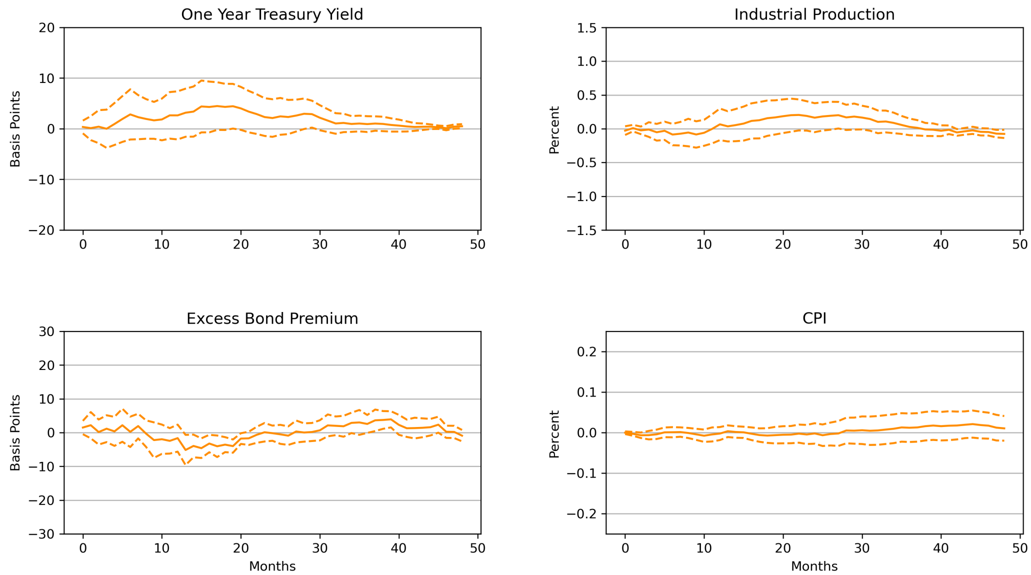
i, and the number of months in the future, indexed by h, such that:

$$Y_{i,t+h} = \theta_{i,k,h} Shock_{k,t} + controls + \eta_{i,k,h} \tag{7}$$

Standard errors are calculated Newey and West (1987) to account for serial correlation of the error terms.

The impulse response functions are responses of macroeconomic variables to a 100 basis point increase to the monetary policy shock. For all the shock series, this represents a contractionary shock. For the Gertler and Karadi (2015) shock this means three-month-ahead FFF price decreases by 100 basis points, so expectations increase by 100 basis points. For the text shock, a 100 basis point increase represents an 100 basis point increase in FFR expectations caused by FOMC announcement, after controlling for the private information of the Fed.

Figure 10: Impulse Responses to 3-Month-Ahead FFF Price Change as Shock



Note: Here the change in the 3-Month-Ahead FFF Price is the monetary policy shock. This is what Gertler and Karadi (2015) use in their shock series. Impulse responses are calculated using the local projection method from Jordà (2005). Confidence bands are at the 90% level. Standard errors are Newey and West (1987) standard errors that correct for heteroskedasticity and autocorrelation. The above are responses to a 100 basis point increase in the shock series (not pictured).

Figure 9 shows that macroeconomic variable responses are much larger for impulses to the cleaned text shock compared to the GK shocks in Figure 10. In particular, a 100 basis point increase to the cleaned text shock is associated with a 80 percentage point decrease in output, a 10 percentage point decrease in inflation, and almost a 20 basis point increase in the excess bond premium after about 10 months. These responses are qualitatively consistent with monetary theory that indicates a contractionary shock should decrease output and inflation.

Figure 10 shows similar results as in Ramey (2016). In particular an increase to the three-month-ahead FFF price, the GK shock, actually produces small increases in output and inflation and a decrease in the excess bond premium. The impulse responses estimated with local projection for the uncleaned text shock, and for the first principal component of FFF price changes are all included in the appendix in Figure C4 and Figure C5, respectively.

One panel in Figure 9 that does not seem to fit is the response of the one year treasury yield to a contractionary text shock. This poor fit is expected considering the correlation between the cleaned text shock and changes in the one year treasury yield has large standard errors and is not statistically different from zero at higher confidence levels.

Overall, these figures show that the cleaned text shock series is representing information different from other shock series in the literature and produces impulse responses that are consistent with the literature.

6 Conclusion

This paper uses a state-of-the-art text analysis neural network to map FOMC statement text to federal funds rate expectations. Using the trained neural network and alternative versions of statements from FOMC meeting materials, I produce a new monetary policy shock series - which I call "cleaned text shocks." An FOMC statement is said to be a large policy shock the neural network predicts it to have a large affect on fed funds futures prices that are above and beyond the average predicted effects of alternative wordings of the statements. In other words, if this shocks series picks up the forward guidance effect of FOMC statements through their word choice.

This paper then compares and contrasts the cleaned text shock series to other monetary policy shock series identified with high-frequency fed funds futures price changes. In terms of summary statistics, the cleaned text shocks are similar to other series. Furthermore, they are similarly correlated with nominal interest rates.

However, I find that the coefficients relating the cleaned text shocks and real interest rates are twice the size of coefficients for other shock series from the literature. This means that shock measures that only use asset price changes are missing information about the effect of monetary policy on the real economy. Differences continue into the impulse response estimation. Responses of output, inflation, nominal interest rates, and the excess bond premium to impulses in Gertler and Karadi (2015) shocks with the cleaned text shocks are dramatically different. As Ramey (2016), using the local projection method to graph impulse responses show the responses of macroeconomic variables are generally not statistically different from zero. Also, qualitatively, they respond in directions that counter the conventional monetary policy theory. However, in response to an increase in the cleaned text shock, macroeconomic variables change as the theory would predict. That is, a contractionary monetary (text) shock produces lower output, lower inflation, and increases to the excess bond premium. Ultimately, this paper shows that monetary policy does influence the economy. Furthermore, it that the Fed affects the economy with its influence over market expectations of future monetary policy and that forward guidance matters.

References

Acosta, M., 2022. The perceived causes of monetary policy surprises. Working Paper URL: https://www.acostamiguel.com/papers/acosta_jmp.pdf.

Ai, H., Bansal, R., 2018. Risk Preferences and the Macroeconomic Announcement Premium. Econometrica 86, 1383–1430. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA14607, doi:10.3982/ECTA14607.

Athey, S., Imbens, G., 2019. Machine Learning Methods Economists Should Know About. Working Paper URL: http://arxiv.org/abs/1903.10075. arXiv: 1903.10075.

Board of Governors of the Federal Reserve System, . The Fed - Federal Open Market Committee. URL: https://www.federalreserve.gov/monetarypolicy/fomc.htm.

Bu, C., Rogers, J., Wu, W., 2019. A Unified Measure of Fed Monetary Policy Shocks. Finance and Economics Discussion Series URL: https://www.federalreserve.gov/econres/feds/a-unified-measure-of-fed-monetary-policy-shocks.htm, doi:https://doi.org/10.17016/FEDS.2019.043r1.

Campbell, J., Evans, C.L., Fisher, J., Justiniano, A., 2012. Macroeconomic Effects of Federal Reserve Forward Guidance. Brookings Papers on Economic Activity 43, 1–80. URL: https://econpapers.repec.org/article/binbpeajo/v_3a43_3ay_3a2012_3ai_3a2012-01_3ap_3a1-80.htm.

Cieslak, A., Schrimpf, A., 2019. Non-monetary news in central bank communication. Journal of International Economics 118, 293–315. URL: http://www.sciencedirect.com/science/article/pii/S002219961830268X, doi:10.1016/j.jinteco.2019.01.012.

Doh, T., Song, D., Yang, S.K., 2020. Deciphering Federal Reserve Communication via Text Analysis of Alternative FOMC Statements. The Federal Reserve Bank of Kansas City Research Working Papers URL: https://www.kansascityfed.org/publications/research/rwp/articles/2020/deciphering-federal-reserve-communication-via-text-analysis, doi:10.18651/RWP2020-14.

Gentzkow, M., Kelly, B., Taddy, M., 2019. Text as Data. Journal of Economic Literature 57, 535–574. URL: https://pubs.aeaweb.org/doi/10.1257/jel.20181020, doi:10.1257/jel.20181020.

Gertler, M., Karadi, P., 2015. Monetary Policy Surprises, Credit Costs, and Economic Activity. American Economic Journal: Macroeconomics 7, 44–76. URL: http://pubs.aeaweb.org/doi/10.1257/mac.20130329, doi:10.1257/mac.20130329.

Gilchrist, S., Zakrajsek, E., 2012. Credit Spreads and Business Cycle Fluctuations. American Economic Review 102, 1692–1720. URL: https://pubs.aeaweb.org/doi/10.1257/aer.102.4.1692, doi:10.1257/aer.102.4.1692.

Gurkaynak, R.S., Sack, B., Swanson, E., 2004. Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements. Finance and Economic Discussion Series , 43.

Gurkaynak, R.S., Sack, B., Wright, J.H., 2007. The U.S. Treasury yield curve: 1961 to the present. Journal of Monetary Economics 54, 2291–2304. URL: https://www.sciencedirect.com/science/article/pii/S0304393207000840, doi:10.1016/j.jmoneco.2007.06.029.

Gürkaynak, R.S., Sack, B., Wright, J.H., 2010. The TIPS Yield Curve and Inflation Compensation. American Economic Journal: Macroeconomics 2, 70–92. URL: https://pubs.aeaweb.org/doi/10.1257/mac.2.1.70, doi:10.1257/mac.2.1.70.

Handlan, A., 2020. Fedspeak matters: Fomc Statements and Monetary Policy Expectations. Working Paper URL: https://sites.google.com/umn.edu/amy-handlan/research.

Hansen, S., McMahon, M., 2016. Shocking language: Understanding the macroeconomic effects of central bank communication. Journal of International Economics 99, S114–S133. URL: http://www.sciencedirect.com/science/article/pii/S0022199615001828, doi:10.1016/j.jinteco.2015.12.008.

Husted, L., Rogers, J., Sun, B., 2017. Monetary Policy Uncertainty. International Finance Discussion Paper 2017, 1–56. URL: https://www.federalreserve.gov/econres/ifdp/files/ifdp1215.pdf, doi:10.17016/IFDP.2017.1215.

Jordà, O., 2005. Estimation and Inference of Impulse Responses by Local Projections. The American Economic Review 95, 161–182. URL: http://www.jstor.org/stable/4132675.

Kuttner, K.N., 2001. Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. Journal of Monetary Economics , 22.

Lucca, D.O., Moench, E., 2015. The Pre-FOMC Announcement Drift. The Journal of Finance 70, 329–371. URL: http://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12196, doi:10.1111/jofi.12196.

Lucca, D.O., Trebbi, F., 2009. Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements. Working Paper 15367. National Bureau of Economic Research. URL: http://www.nber.org/papers/w15367, doi:10.3386/w15367.

Nakamura, E., Steinsson, J., 2018. High-Frequency Identification of Monetary Non-Neutrality: The Information Effect. The Quarterly Journal of Economics 133, 1283–1330. URL: https://academic.oup.com/qje/article/133/3/1283/4828341, doi:10.1093/qje/qjy004.

Neuhierl, A., Weber, M., 2018. Monetary Momentum. Working Paper 24748. National Bureau of Economic Research. URL: http://www.nber.org/papers/w24748, doi:10.3386/w24748.

Newey, W.K., West, K.D., 1987. Hypothesis Testing with Efficient Method of Moments Estimation. International Economic Review 28, 777–787. URL: https://www.jstor.org/stable/2526578, doi:10.2307/2526578.

Ramey, V.A., 2016. Chapter 2 - Macroeconomic Shocks and Their Propagation, in: Taylor, J.B., Uhlig, H. (Eds.), Handbook of Macroeconomics. Elsevier. volume 2, pp. 71–162. URL: http://www.sciencedirect.com/science/article/pii/S1574004816000045, doi:10.1016/bs.hesmac.2016.03.003.

Romer, C.D., Romer, D.H., 2004. A New Measure of Monetary Shocks: Derivation and Implications. American Economic Review 94, 43.

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., Le, Q.V., 2020. XLNet: Generalized Autoregressive Pretraining for Language Understanding. Working Paper URL: http://arxiv.org/abs/1906.08237. arXiv: 1906.08237.

A Overview of Training Algorithm

1. Fix the collection of text (call this "corpus 1")
2. Prepare the text to be an numerical input to the neural network
   (a) Break words into sub-word units (called tokens)
   (b) Create 768x1 vector for each token based on co-occurrence of sub-word units in the corpus (a clustering algorithm to train the vector values so that similar words have similarly oriented vectors in 768-dimensional space)
   (c) Add special tokens to indicate ends of sentence and an observation level identifier
   (d) Add padding to make all text inputs the same length (256 for now, robustness with 512 and 900 later)
3. Train the neural network for task 1 on corpus 1
   (a) Fix the network structure and the hyperparameters (ie learning rate)
   (b) Update parameters in the network to increase prediction accuracy for training data (predicting missing words from text inputs)
   (c) Stop updating parameters
   (d) Evaluate neural network: prediction accuracy on data not used for training (testing data)
   (e) Go back to initial step and restructure neural network if needed
4. Fine-tune neural network for task 2 on corpus 2
   (a) Add additional layer to network to handle new task
   (b) Update parameters to increase prediction accuracy for new training data
   (c) Stop updating
   (d) Evaluate neural network: prediction accuracy on data not used for training (testing data)

B Table Appendix

B.1 Summary Statistics

Table B1: Statistics of Monetary Shocks, FOMC Meetings from 2005-2014

	PC1 FFF Prices	Text Shock	Cleaned Text Shock	NS Shock	GK Shock: FF4	GK shock: TY1(FF4)
count	80	80	80	74	80	80
mean	-0.0000	-0.0027	0.0011	0.0039	-0.0018	-0.0042
std	0.0215	0.0158	0.0113	0.0321	0.0395	0.0294
min	-0.1009	-0.0900	-0.0685	-0.1452	-0.19	-0.1441
25%	-0.0012	-0.0058	-0.0029	-0.0034	-0.005	-0.0066
50%	0.0013	-0.0007	0.0022	0.0076	0	-0.0029
75%	0.0027	0.0031	0.0060	0.0186	0.0063	0.0017
max	0.0631	0.0675	0.0406	0.0679	0.115	0.0825

Note: "NS shock" is from Nakamura and Steinsson (2018) and is the first principal component of fed funds futures and Eurodollar futures. "PC1 FFF Prices" is the first principal component of fed funds futures prices representing target fed funds rate expectations at the current and next FOMC meetings. This is the NS shock without Eurodollar futures. "GK Shock: FF4" is the change in the 3 month ahead fed funds future price (FF4)."GK shock: TY1(FF4)" is the daily change in the 1 year treasury yield instrumented with the FF4. Work in the text is for the latter version of the GK shock.

Table B2: Statistics of Nominal Interest Rate Changes, FOMC Meetings, 2005-2014

	$\Delta\ TY_1$	$\Delta\ TY_2$	$\Delta\ TY_3$	$\Delta\ TY_5$	$\Delta\ TY_{10}$
count	80	80	80	80	80
mean	-0.0009	0.0018	0.0025	0.0012	0.0004
std	0.0544	0.0661	0.0772	0.0918	0.0923
min	-0.2045	-0.2641	-0.3477	-0.4708	-0.5189
25%	-0.0198	-0.027	-0.0314	-0.0385	-0.0356
50%	0.0019	-0.0008	0.0009	0.008	0.0135
75%	0.0189	0.0322	0.0469	0.0444	0.0569
max	0.2023	0.2296	0.2263	0.1844	0.2019

Note: The above represent the daily change in the h-year treasury yields (ΔTY_h). The yield change is evaluated from end-of-day before FOMC announcement day to the end of the day of the FOMC announcement. Data is from Gurkaynak et al. (2007).

Table B3: Statistics of Real Interest Rate Changes, FOMC Meetings, 2005-2014

	$\Delta\ TIPS_2$	$\Delta\ TIPS_3$	$\Delta\ TIPS_5$	$\Delta\ TIPS_{10}$
count	80	80	80	80
mean	-0.0072	-0.0081	-0.0074	-0.0047
std	0.1183	0.1141	0.1094	0.0963
min	-0.5215	-0.5499	-0.5818	-0.5705
25%	-0.0467	-0.0476	-0.0509	-0.0353
50%	-0.0024	0.0032	0.009	0.0072
75%	0.0484	0.0522	0.0451	0.0463
max	0.3637	0.2998	0.2187	0.1569

Note: The above represent the daily change in the h-year TIPS yields ($\Delta TIPS_h$). The yield change is evaluated from end-of-day before FOMC announcement day to the end of the day of the FOMC announcement. Data is from Gürkaynak et al. (2010).

Table B4: Statistics for Impulse Response variables, Monthly for 2005-2014

	log IP	log CPI	EBP	TY_1
count	120	120	120	120
mean	4.60	5.39	0.04	1.64
std	0.05	0.05	0.85	1.88
min	4.47	5.29	-0.92	0.09
25%	4.57	5.35	-0.40	0.20
50%	4.61	5.40	-0.22	0.42
75%	4.63	5.44	-0.01	3.41
max	4.67	5.48	3.47	5.20

Note: All logs are natural logarithms. Industrial production (IP) and Consumer Price Index (CPI) are sourced from FRED. The Excess Bond Premium (EBP) is from Gilchrist and Zakrajsek (2012) and here is in percentage points. The 1 year Treasury Yield (TY_1) is from Gurkaynak et al. (2007).

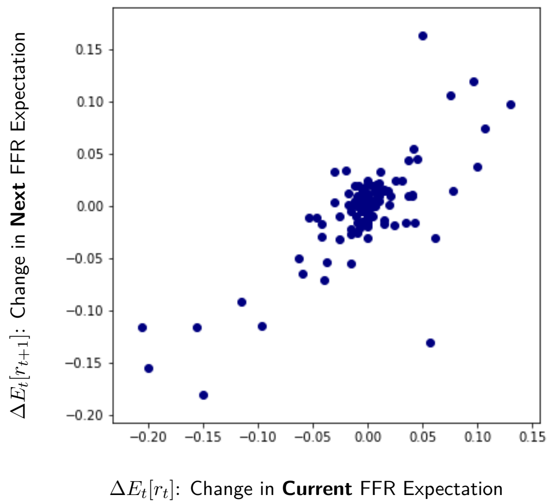
Table B5: Statistics for Monetary Shocks, Monthly 2005-2014

	Text Shock	Cleaned Text Shock	PC1 FFF	GK Shock: FF4	GK Shock: FF4 rolling average
count	120	120	120	120	90
mean	-0.0018	0.0007	-0.0000	-0.0012	-0.005371
std	0.0129	0.0092	0.0175	0.0322	0.032843
min	-0.09	-0.0685	-0.1009	-0.1900	-0.206291
25%	-0.0016	-0.0014	0	0	-0.0048
50%	0	0	0	0	0
75%	0.0008	0.0036	0.0013	0	0.0037
max	0.0675	0.0406	0.0631	0.1150	0.0561

Note: For all but the last column, the shocks are zero for any month that does *not* have an FOMC meeting. "GK Shock: FF4 rolling average" is aggregated as a rolling average over the past month to create the monthly series. For this latter column, it means that months without FOMC meetings will have non-zero shock values.

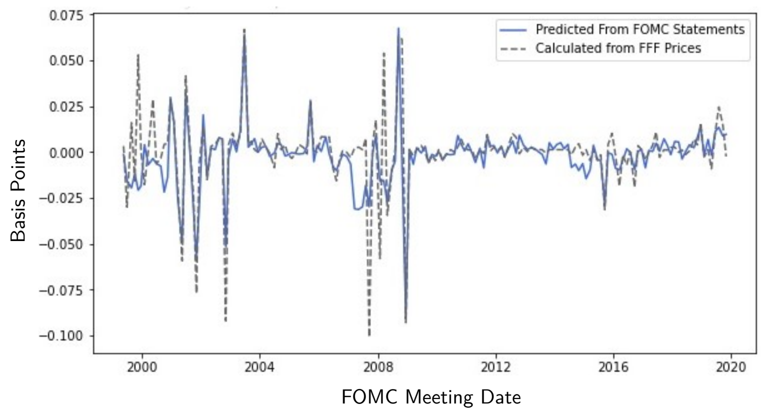
C Graph Appendix

Figure C1: Change in Expectations of Federal Funds Rate (FFR)



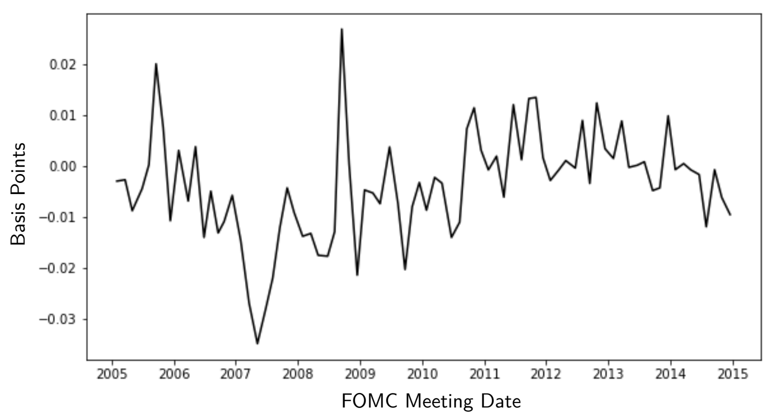
Note: The change in expectations today of the federal funds rate at the current meeting $\Delta E_t[r_t]$ and the next meeting $\Delta E_t[r_{t+1}]$ are calculated from changes in fed funds futures prices.

Figure C2: Change in Expectations from FFF Prices vs. Text Prediction



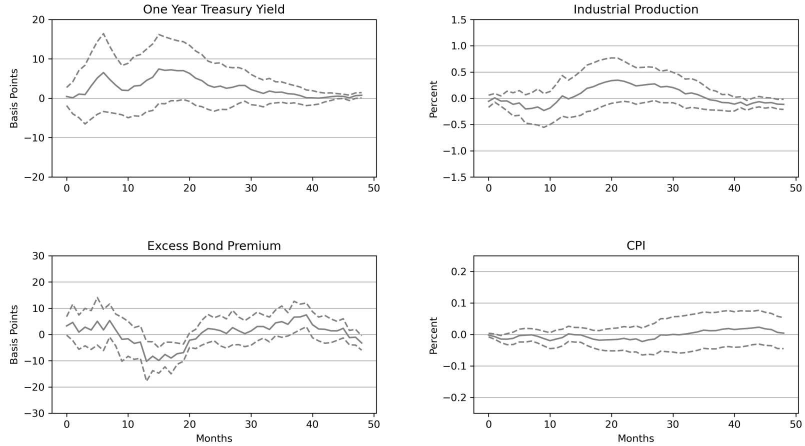
Note: The gray, dashed line is the change in expectations calculated from fed funds futures (FFF) prices. It is the first principal component of two variables: changes in expectations of the federal funds rate for the current meeting and the next meeting. These expectations are calculated from changes in FFF prices from 10 minutes before to 20 minutes after the FOMC announcement is released. The blue, solid line, $\widehat{\Delta E[r]}_{text}$, is the prediction of the previous variable from the FOMC statement text and the neural network.

Figure C3: Average Predicted $\Delta E[r]$ across Alternative Statements



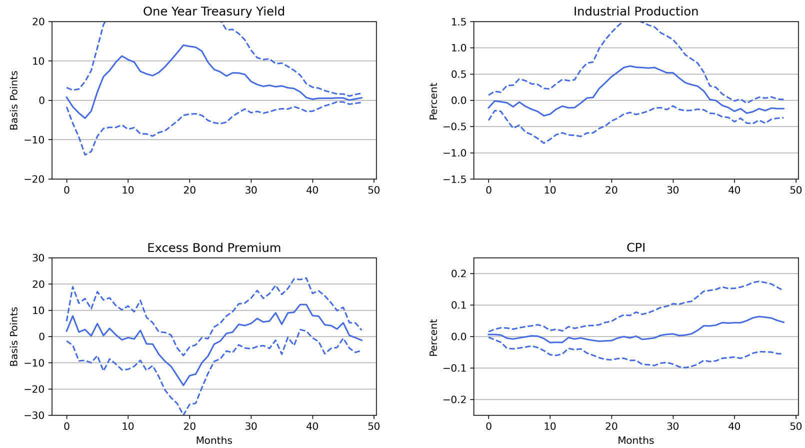
Note: Alternative versions of FOMC statements are included in the Bluebooks and Tealbooks (FOMC meeting materials). I feed each alternative into the trained neural network to get a predicted change in FFR expectations. This graph is then the average of predictions from every alternative for each FOMC meeting from 2005-2014.

Figure C4: Impulse Responses to First Principal Component of FFF, $\Delta E[r]_{FFF}$



Note: Impulse responses are calculated using the local projection method from Jordà (2005). Confidence bands are at the 90% level. Standard errors are Newey and West (1987) standard errors that correct for heteroskedasticity and autocorrelation. The above are responses to a 100 basis point increase in the shock series (not pictured).

Figure C5: Impulse Responses to Text Shock, $\widehat{\Delta E[r]}_{Text}$



Note: Impulse responses are calculated using the local projection method from Jordà (2005). Confidence bands are at the 90% level. Standard errors are Newey and West (1987) standard errors that correct for heteroskedasticity and autocorrelation. The above are responses to a 100 basis point increase in the shock series (not pictured).