How Much Information Do Monetary Policy Committees Disclose? Evidence from the FOMC Minutes and Transcripts

Mikael Apel†, Marianna Blix Grimaldi‡, and Isaiah Hull†

†Sveriges Riksbank
‡Swedish National Debt Office

1. Abstract

The purpose of central bank minutes is to give an account of monetary policy meeting discussions to outside observers, thereby enabling them to draw informed conclusions about future policy. However, minutes are by necessity a shortened and edited representation of a broader discussion. Consequently, they may omit information that is predictive of future policy decisions. To investigate this, we compare the information content of the FOMC's minutes and transcripts, focusing on three dimensions which are likely to be excluded from the minutes: 1) the committee's degree of hawkishness; 2) the chairperson's degree of hawkishness; and 3) the level of agreement between committee members. We measure committee and chairperson hawkishness using a novel dictionary that is constructed using the FOMC's minutes and transcripts. Agreement is measured by performing deep transfer learning, a technique that involves training a deep learning model on one set of documents – U.S. congressional debates – and then making predictions on another – FOMC transcripts. Our findings suggest that transcripts are more informative than minutes and heightened committee agreement typically precedes policy rate increases.

2. Measurement

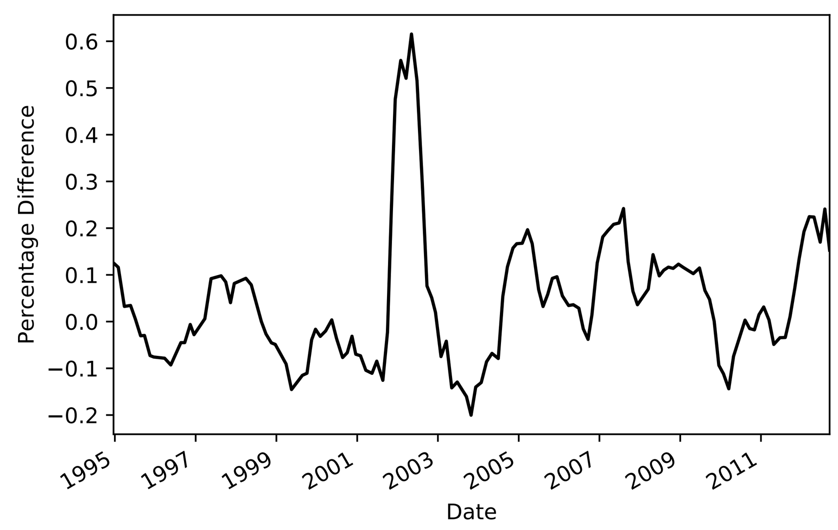
We introduce two text-based measures of policy inclination. The first captures net hawkishness using a novel dictionary-based method, which builds on the approach in Apel and Blix Grimaldi (2012). This revised version of the dictionary is partitioned into three sections, each of which has direct relevance for part of the Fed's dual mandate: 1) inflation; 2) economic activity; and 3) employment. We then compute the set of most frequently used terms, drawing from the unigrams and bigrams in the minute and transcript texts, and associate those that are policy relevant with a dictionary topic. We next select the modifiers that are most frequently used in the same sentence as each noun to determine whether the term was used in a hawkish or dovish context. For example, for "inflation," we pick modifiers such as "rising" and "accelerating." We apply this dictionary to the FOMC's minutes and transcripts, as well as chairperson's text within the transcripts.

Our measure of agreement is constructed by performing transfer learning. We first train a deep learning model with a long short-term memory architecture to classify agreement in U.S. congressional debate texts. We do this by making use of a pre-labelled dataset that contains information about whether a speaker voted against or in favor of a bill. We then apply the trained model to FOMC transcripts to identify whether each speaker has adopted a tone that is generally agreeable or disagreeable. The two examples given below show cases with high agreeable ratings for the training set – U.S. congressional debates – and the test set – FOMC discussions.

Train Set: *I thank the gentleman for yielding me this time and for his great work on this bill. Mr. chairman, this country needs to create a new energy landscape that begins shrinking our disproportionate reliance on foreign energy sources and begins building one that places American ingenuity, producers and consumers at the forefront. I want to highlight one provision and that is the provision that significantly strengthens the important leaking underground storage tank program. The bill increases state funding... I urge their support and the support of the bill.*

Test Set: *Based on research from my staff, I have also lowered my estimate of ... real gdp growth ... reflecting lower expectations of trend growth for both the labor force and labor productivity.*

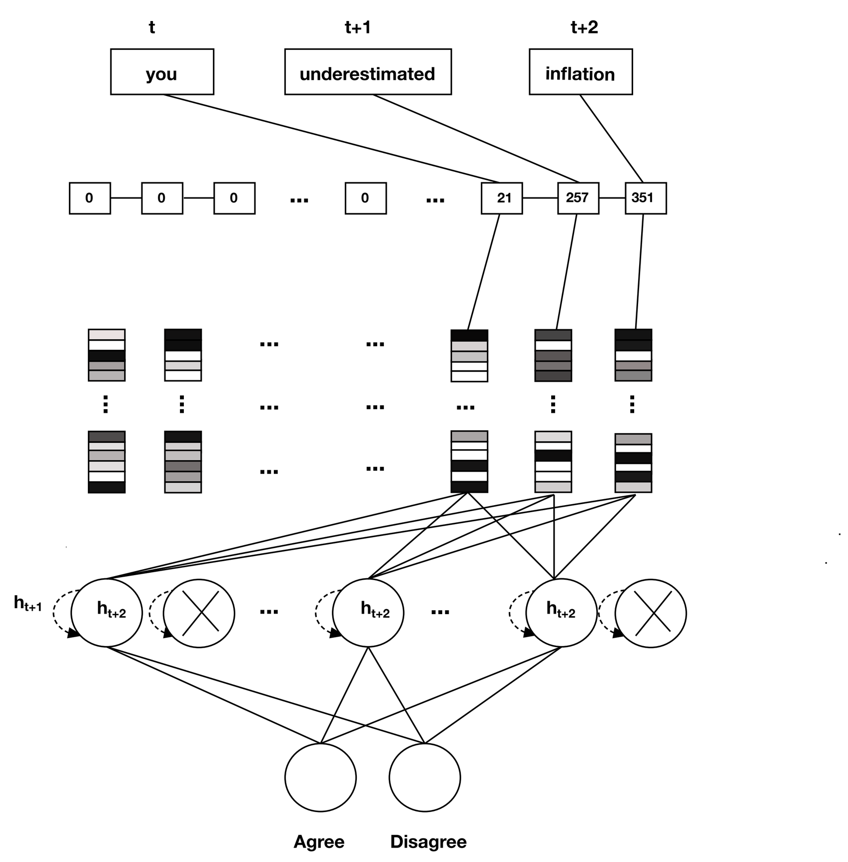
Figure 1: *Rolling Percentage Difference Between Transcript and Minute Index Levels*



Notes: The plot above shows the rolling average of the percentage difference between transcripts and minute index levels. The underlying series contain one observation per FOMC meeting. There are typically eight FOMC meetings per year with an interval of 5 to 8 weeks between each.

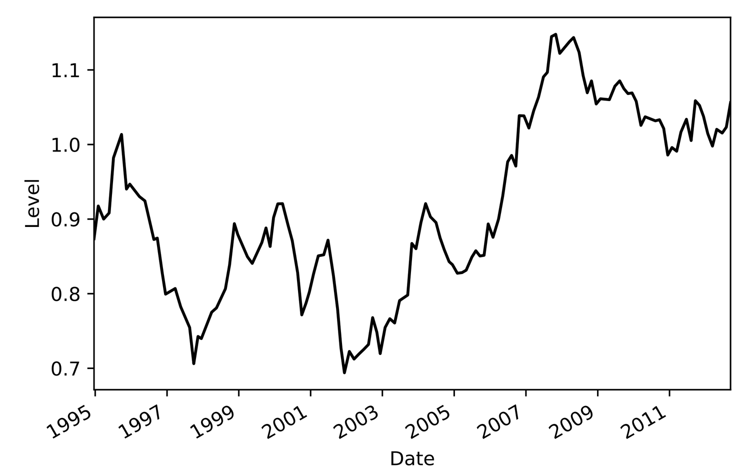
Agreement

Figure 2: *Long Short-Term Memory Model Architecture*



Notes: The figure above shows the architecture for a long short-term memory model (LSTM), which we use to perform transfer learning. The model first takes words represented by sparse vectors as inputs. It translates these into integers, which are then connected to word embeddings. The word embeddings are then processed by long short-term memory cells and passed to a dropout layer. Finally, the output of the dropout layer is flattened into a vector and connected to the output layer via a dense layer. The model's output is the probability that the input sequence contains dialogue in which the speaker is agreeing, rather than disagreeing.

Figure 3: *Rolling Measure of Agreement*



Notes: The plot above shows the 6-month rolling average of FOMC member agreement, measured using transcripts. We construct a deep learning model to predict agreement in U.S. congressional debates. We then use the model to predict agreement in FOMC transcripts.

3. Empirical Specification

Our main empirical specification consists of an ordinal logit where the dependent variable, $Policy_t$, can take on three values: -1 for a monetary easing; 0 for no change; and +1 for a monetary tightening. $Transcripts_{t-1}$ is the level of the net hawkishness index for the transcripts, $Minutes_{t-1}$ is the level of the net hawkishness index for the minutes, $Agreement_{t-1}$ is a measure of agreement among committee members, $Chair_{t-1}$ is a measure of the chairperson's net hawkishness, and X_{t-1} is a vector of controls.

$$\begin{aligned} Policy_t = \alpha + \beta_1 Transcripts_{t-1} + \beta_2 Minutes_{t-1} \\ + \beta_4 Agreement_{t-1} + \beta_5 Chair_{t-1} + \beta_6 X_{t-1} \end{aligned} \tag{1}$$

4. Results

Table 1: *Ordered probit marginal effects: Predictive value of central bank minutes and transcripts for policy decisions.*

	(1) (Net)	(2) (Net)	(3) (Net)	(4) (Hawk)
Policy Decision = Ease.				
Minutes	-0.04	0.05	-0.03	-0.08
	(0.055)	(0.063)	(0.578)	(0.11)
Transcripts	-0.37***	-0.28***	-0.42***	-0.74***
	(0.038)	(0.107)	(0.136)	(0.25)
Chair	0.02	0.04	0.03	-0.04
	(0.038)	(0.042)	(0.040)	(0.076))
Agreement	-0.15*	-0.15*	-0.18*	-0.169*
	(0.087)	(0.091)	(0.091)	(0.087)
Policy Decision = No change.				
Minutes	-0.01	0.01	-0.01	-0.03
	(0.026)	(0.024)	(0.018)	(0.051)
Transcripts	-0.14	-0.07	-0.08	-0.27
	(0.165)	(0.113)	(0.175)	(0.033)
Chair	0.00	-0.01	0.01	-0.02
	(0.017)	(0.170)	(0.010)	(0.034)
Agreement	-0.06	-0.04	0.03	-0.06
	(0.078)	(0.064)	(0.076)	(0.078)
Policy Decision = Tighten.				
Minutes	0.05	-0.06	0.07	0.11
	(0.074)	(0.078)	(0.068)	(0.149)
Transcripts	0.51***	0.34***	0.50***	1.01***
	(0.158)	(0.130)	(0.160)	(0.316)
Chair	-0.03	0.04	-0.03	0.059
	(0.052)	(0.050)	(0.050)	(0.10)
Agreement	0.23**	0.19*	0.21**	0.23**
	(0.052)	(0.110)	(0.110)	(0.042)
Controls	YES	YES	YES	YES
TFFR	YES	NO	YES	YES
Δ TFFR	NO	YES	NO	NO
QE	NO	NO	YES	NO
Log-Likelihood	-75.11	-73.52	-75.20	-75.11
Pseudo-R^2	0.38	0.41	0.38	0.38
N	136	135	136	136

Notes: All specifications use an ordered probit. The dependent variable in columns (1), (2), and (4) is an ordinal variable that can take on one of three values: -1 when the target federal funds is lowered; 0 when the rate remains unchanged; and +1 when it rises. In column (3), we treat quantitative easing (QE) as expansionary policy and use -1 to indicate either QE or a rate reduction. The set of controls includes inflation, non-farm payroll growth, house price growth, stock price growth, and inflation expectations. We also control for the lagged target federal funds rate target in specifications (1), (3), and (4) and the first difference of the lagged federal funds rate target in (2). * $p < .1$, ** $p < .05$, *** $p < .01$.

5. Conclusion

We introduce a novel measure of policy inclination using dictionary-based methods, which captures hawkishness in central bank texts, as well as a measure of agreement, which is constructed using transfer learning. We find that the transcripts contain sentiment content that predicts future policy decisions, even after controlling for the sentiment of minutes, as well as macroeconomic and financial variables. We also find that heightened committee agreement appears to precede monetary tightening, suggesting that committees may be more reluctant to raise rates and face negative media coverage until they have achieved an internal consensus.