The Voice of Monetary Policy

Gorodnichenko, Pham, and Talavera

6th March 2024

Tone vs Words

Have you ever inquired about someone's well-being, received an affirmative response, and yet found yourself doubting their sincerity? What do you think prompts that disbelief?

Mehrabian (1971): 7-38-55

The breakdown of meaning:

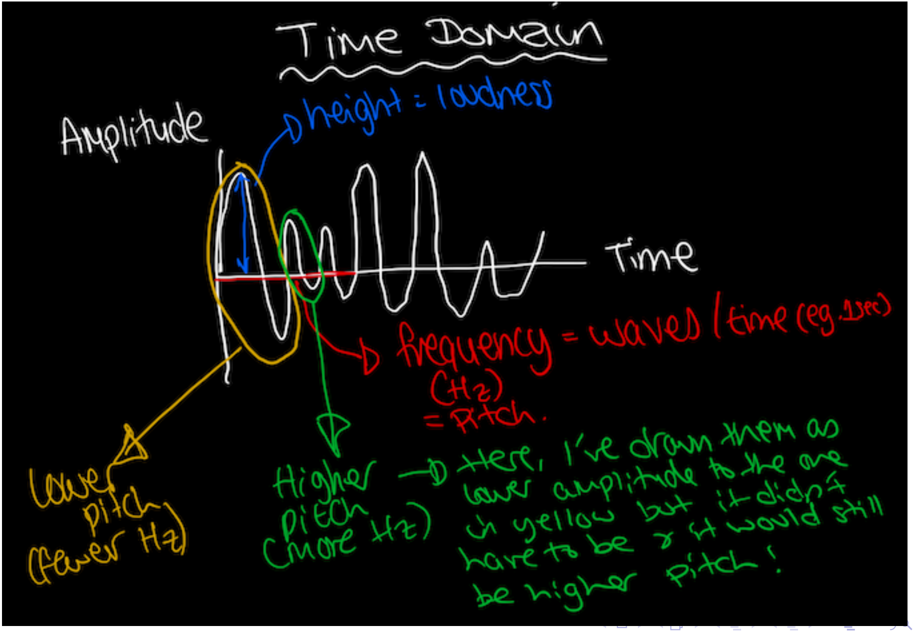
- 7% is words
- 38% is tone
- 55% is non-verbal

What's the added value?

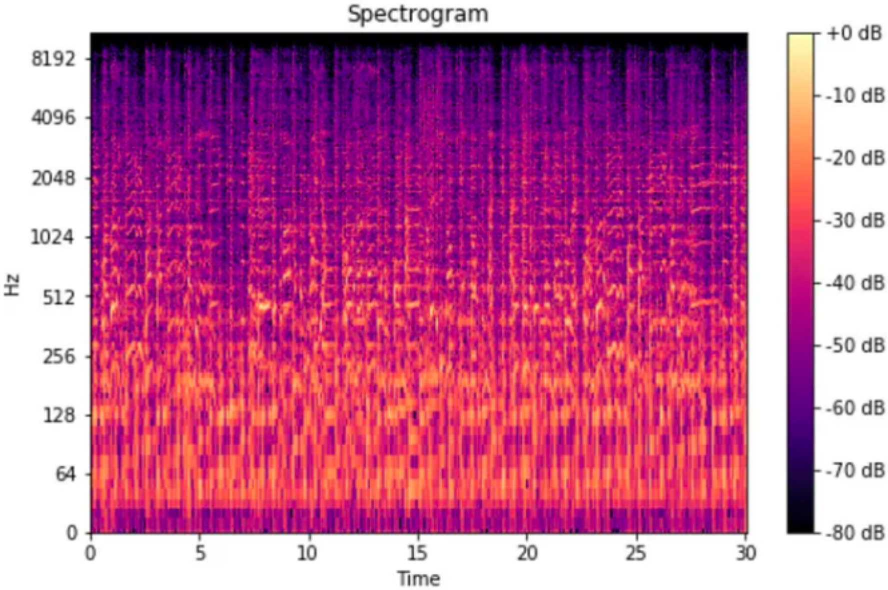
Numerous studies have explored the impact of central bankers' language on financial markets, yet this paper is the first to explore the vocal tone channel. Instead of solely focussing on textual analysis, it examines the way messages are delivered vocally, investigating whether the tone of voice can influence the financial markets.

Curti and Kazinnik (2021) is the most similiar in terms of work. They analyse the response of the financial market to variations in the chair's facial expressions during post-FOMC conferences. However, just focussing on voice is potentially a bit more robust as camera angles etc. come into play when watching a video.

Sound Wave Lesson



Lesson cont.


Spectrogram
8192
4096
2048
1024
512
256
128
64
0
Hz
0
5
10
15
20
25
30
Time
+0 dB
-10 dB
-20 dB
-30 dB
-40 dB
-50 dB
-60 dB
-70 dB
-80 dB

Who is Mel?

- It's a scale
- It's a logarithmic transformation of a signal's frequency
- When the frequency domain is **not** converted to the Mel scale, each unit of frequency corresponds to an equal increment of pitch across the spectrum
- BUT humans perceive differences in pitch less distinctly at higher frequencies. Thus, this non-linear scale reflects how we are more sensitive to changes in frequency(/pitch) at lower frequencies (than higher ones)

Thus, if the goal is to understand human auditory perception - like the goal of this paper - the Mel scale should be used.

Training a neural network model

- 2 datasets that attach emotion to audio: RANDESS & TESS
- Use Librosa package in Python to extract useful features of the audio (e.g. Mel & Chroma coefficients)
- Then use Keras to find patterns in the audio data.
- Keras builds a neural network model where the user specifies certain things (e.g. number of layers, neurons etc)
- Each layer is tasked with learning different aspects of the data
- Neurons are like processing units (i.e. it receives the input, performs some computation on that input, and passes its output to the next layer's neurons).

Overfitting

- 80% for training
- 20% for predicting
- A dropout rate of 30%

How is this applied to monetary policy?

- Use press conferences - particularly the Q&A part.
- April 2011 - June 2019: 68 meetings and 36 press conferences.
- Download videos from the Fed's YouTube channel but only uses the audio track.
- 5 emotions: Happy, (pleasantly) surprised, neutral, sad, and angry.

Voice Tone

$$VoiceTone = \frac{Positive\ answers - Negative\ answers}{Positive\ answers + Negative\ answers}$$

Ben Bernanke $>$ Janet Yellen $>$ Jerome Powell

- They also checked other media sources to see if they could trust the results of their model.

Exogenous Variation

One would expect the actual words and how they're said to align with each other. BUT without external variation in how policy messages are conveyed, it's difficult to isolate the impact of tone.

- Solution: Text sentiment analysis.

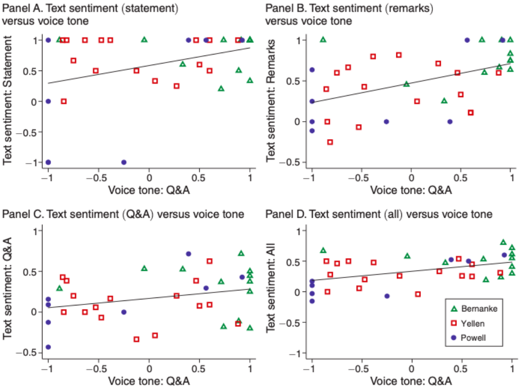


Figure 1. Voice Tone versus Text Sentiment

Note: This figure shows the joint distribution of voice tone and text sentiment across FOMC meetings.

Exogenous Variation cont.

What about being able to isolate the effect of tone from policy actions/stance?

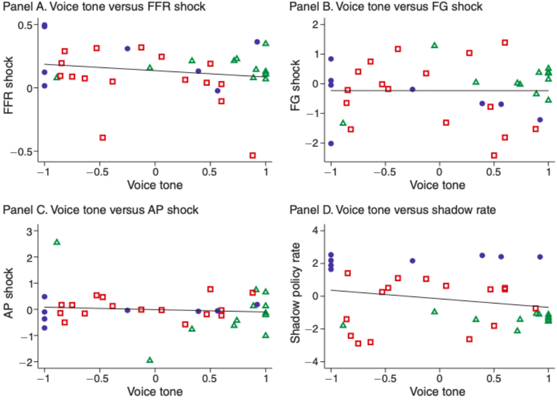


FIGURE 2. POLICY WORDS VERSUS ACTIONS

Notes: This figure shows the joint distribution of voice tone and policy actions/stance. Federal funds rate (FFR), forward guidance (FG), and asset purchase (AP) shocks are from Swanson (2021). The shadow policy rate is from Wu and Xia (2016).

Voice Tone on Financial Indicators: Equation

Follows of the specification of Jordà (2005), by using local projections.

$$
\begin{aligned}
Outcome_{t,t+h} = {} & b_0^{(h)} + b_1^{(h)} VoiceTone_t + b_2^{(h)} TextSentiment_t \\
& + b_3^{(h)} FFRShock_t + b_4^{(h)} FGShock_t + b_5^{(h)} APShock_t \\
& + b_6^{(h)} ShadowRate_t + b_7^{(h)} \mathbf{1}\{NoPressConference_t\} \\
& + error_t^{(h)},
\end{aligned}
$$

Policy shocks and the shadow rate are there to control for the ‘actions’ of the Fed.

Voice Tone on Financial Indicators: Results

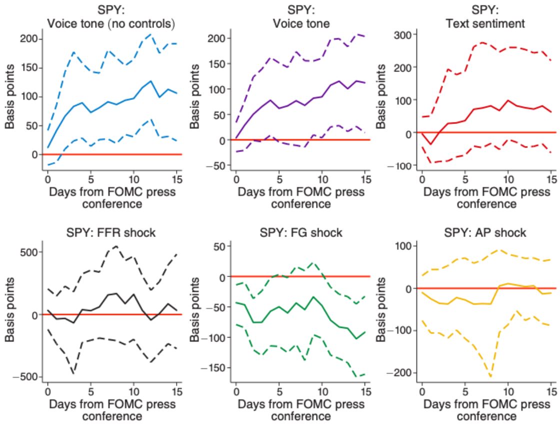


FIGURE 3. RESPONSE OF SPY ETF (S&P 500) TO POLICY ACTIONS AND MESSAGES

Notes: This figure reports the estimated slope coefficients b (specification (4)) for policy communication/actions. Dashed lines show 90 percent bias-corrected and accelerated bootstrap confidence intervals.

A more positive voice tone leads to an increase in share prices. They also go on to find that a more positive tone leads to a decrease in expected inflation.

The Shape of the Response

- Previous literature has shown that financial variables typically tend to respond immediately and sharply to the Fed's action.
- But what we can see from the previous figure, the effects of voice tone and text sentiment on the financial market is not immediate.
- In the paper, they argue that perhaps the tone is leading indicator of subsequent policy communication.

Twitter

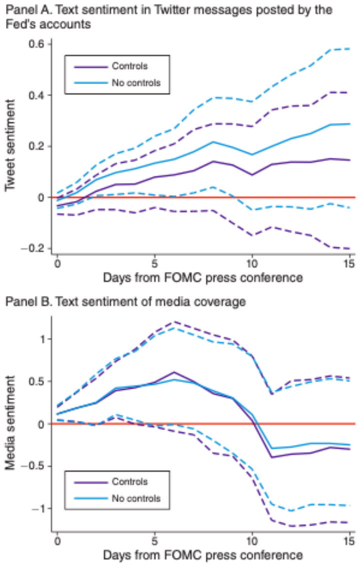


FIGURE 16. POST-FOMC POLICY COMMUNICATION AND MEDIA COVERAGE

Summary

- Uses a machine learning model to rate each piece of audio as either happy, neutral, or sad.
- They find that tone can materially move financial markets.
- Suggests there is more to policy communication than just reading words.