# USING AI TO INVESTIGATE THE EFFECTS OF CENTRAL BANK COMMUNICATIONS: *IT'S NOT JUST WHAT THEY, SAY BUT HOW THEY SAY IT!*

*Michelle Alexopoulos, Department of Economics & Faculty of Information, University of Toronto*

# BASED ON A PROJECT WITH…

Xinfen Han, Bank of Canada

Oleksiy Kryvtsov, Bank of Canada

Xu Zhang,  Bank of Canada

LINK: https://ideas.repec.org/p/tor/tecipa/tecipa-737.html

*Disclaimer: The views expressed are solely those of the researchers and do not necessarily reflect the views of the Bank of Canada.*

# OVERVIEW OF PROJECT

- Examine the importance and impact of the U.S. Federal Reserves' policy communications
- Measure aspects of communication (AI, NLP, etc.)
- Test if (and how) words chosen, body language, & tone of voice effect S&P500 and the VIX

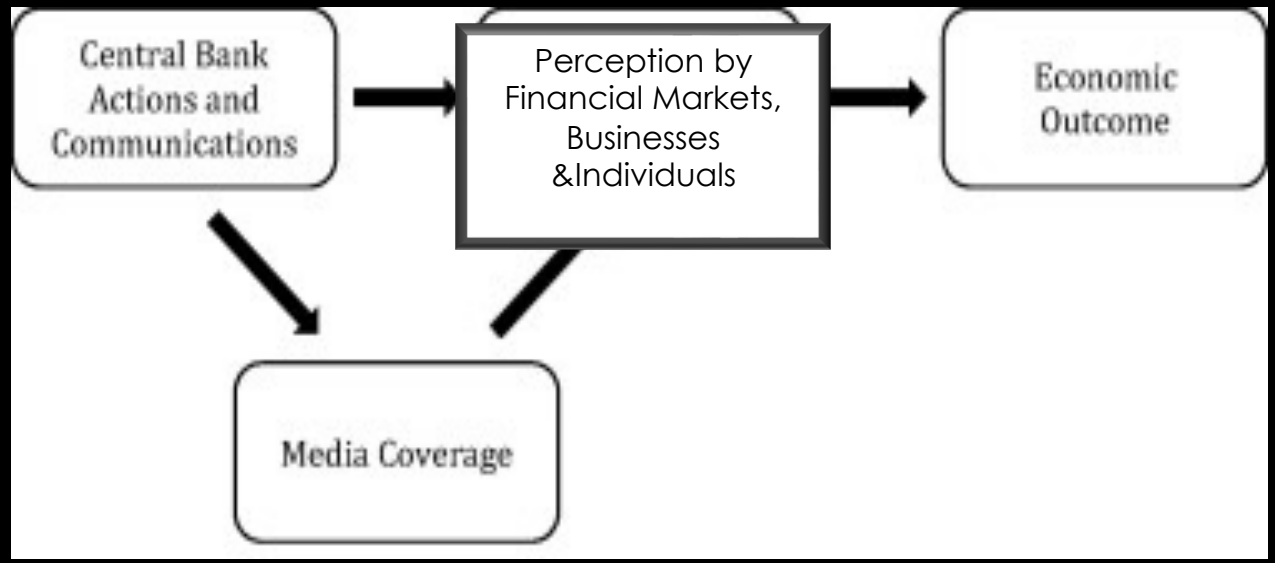
# IMPORTANCE OF POLICY COMMUNICATION

Monetary policies (e.g., interest rate, forward guidance, QE, etc.)

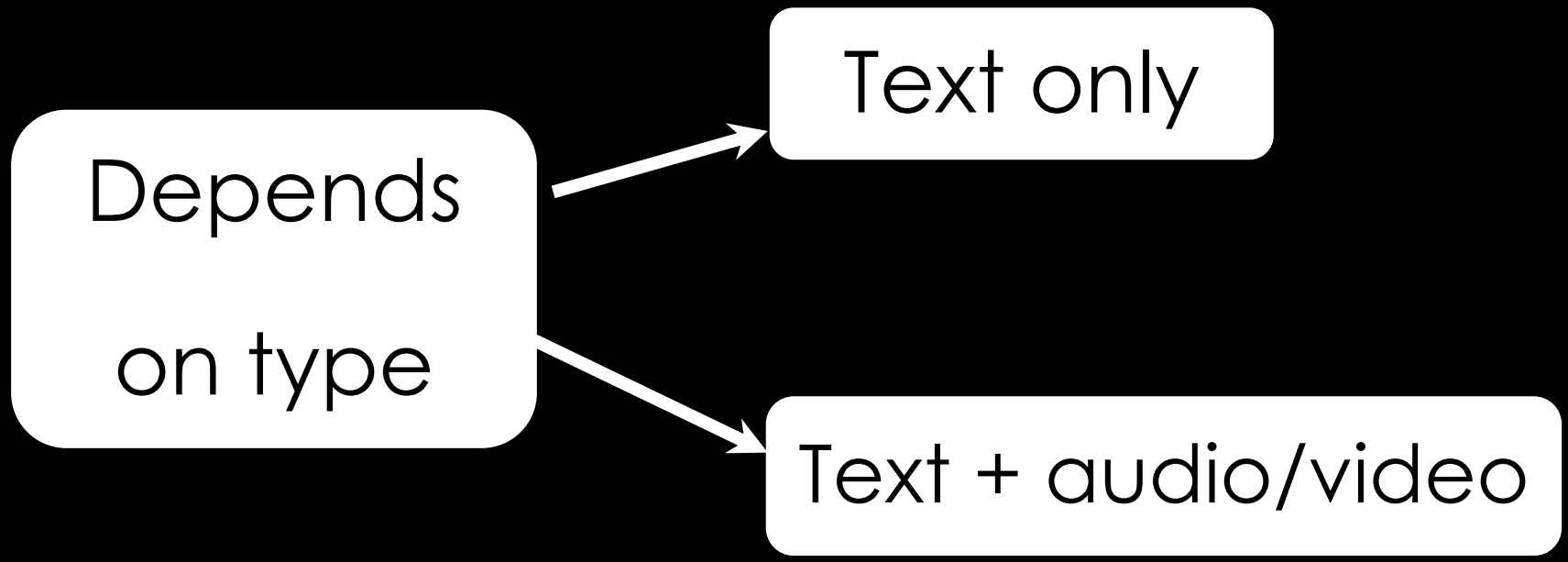
- Have large impacts on the welfare of businesses and individuals
- Are often complex

To Be Effective, Central Bank Communication needs to:

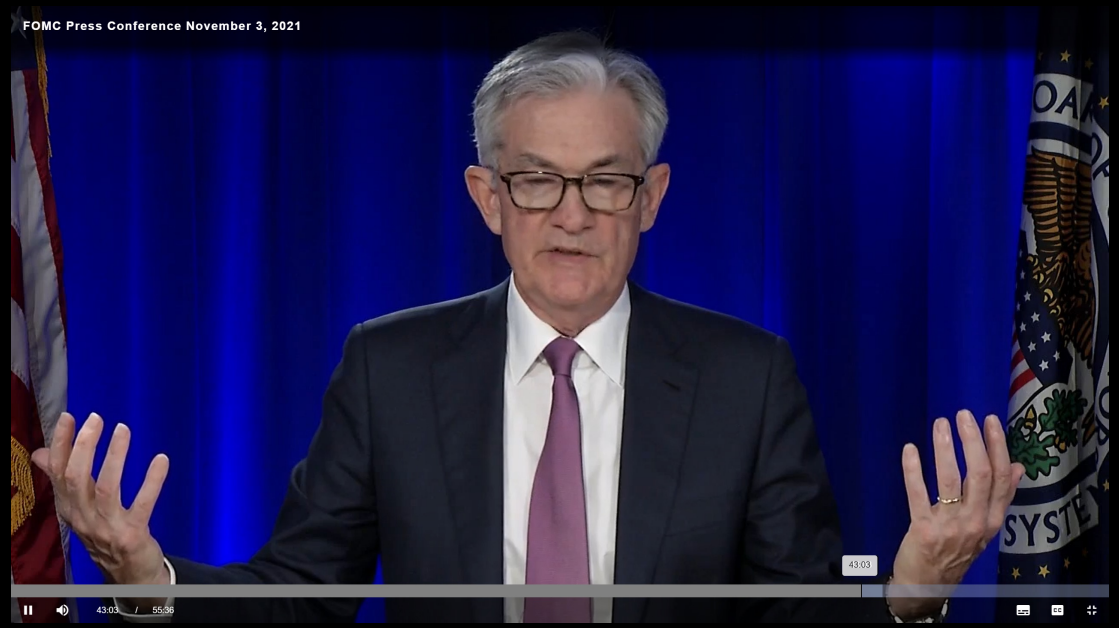
→ Be Accurate
→ Be Clear
→ Be Perceived as Credible
→ Reach target audience



# WHAT MAKES UP POLICY COMMUNICATION?

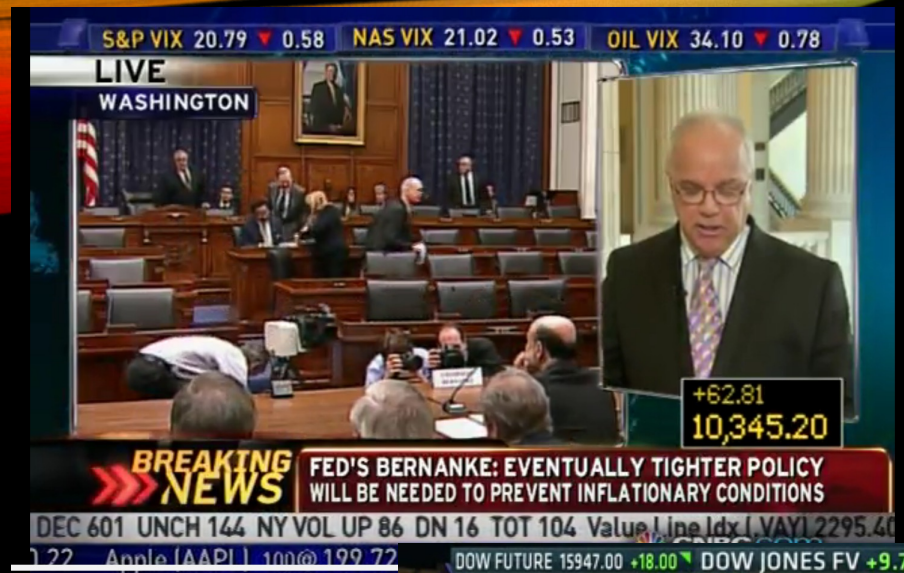


# EXAMPLES OF DIRECT COMMUNICATION



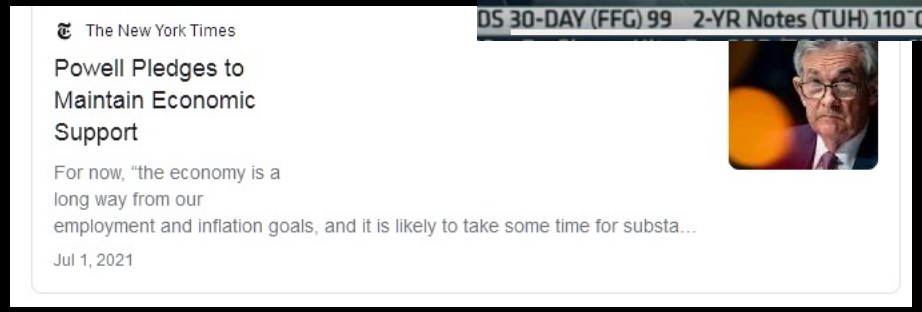

"…the Committee decided to raise the target range for the federal funds rate to 3/4 to 1 percent and anticipates that ongoing increases in the target range will be appropriate…" FOMC statement May 4, 2022 @2pm

# EXAMPLES OF INDIRECT COMMUNICATION

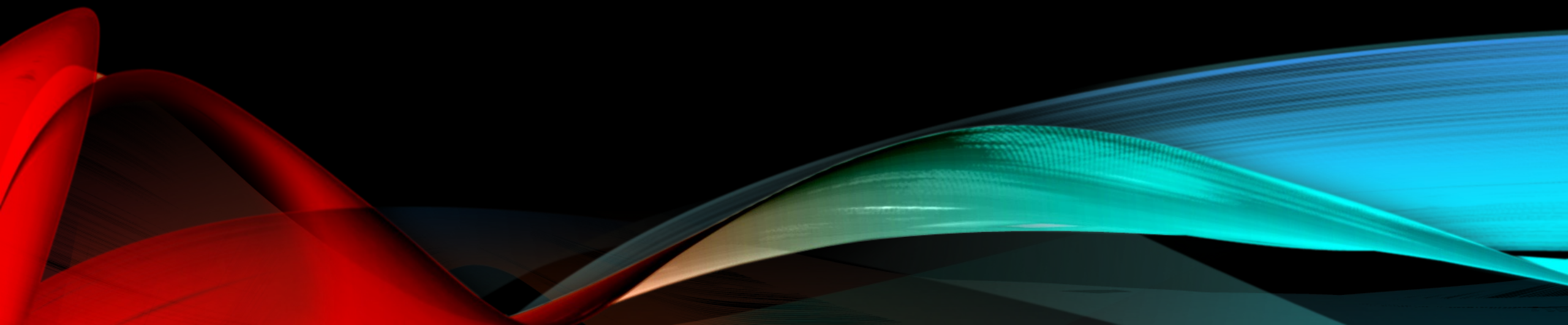


BEN BERNANKE SAYS IT WOULD BE

PANEL AND SIGNALED THERE WOULDN'T BE A MAJOR CHANGE IN

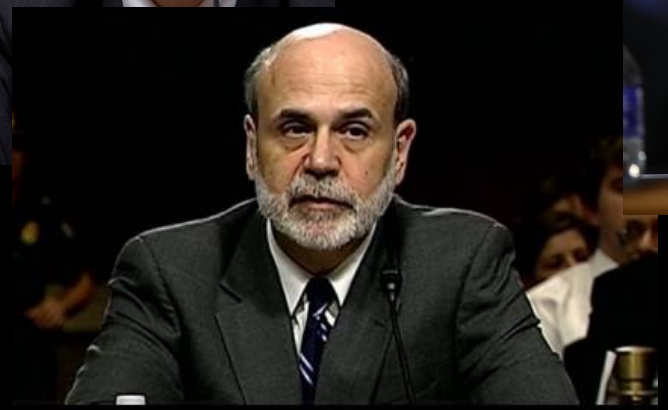


# WHAT FORMS THE MESSAGE?

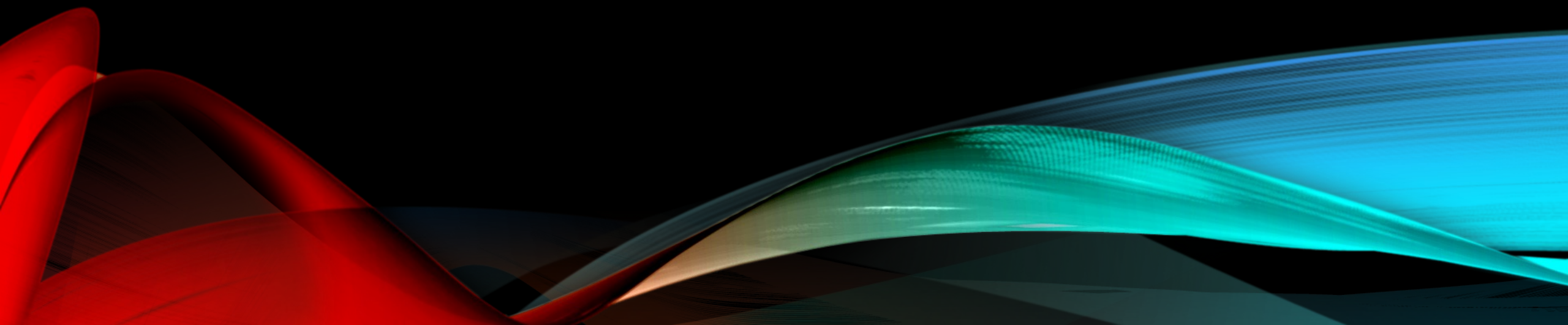


# ELEMENTS OF COMMUNICATION: MORE THAN WORDS

- Have you heard the statement over 90% of communication comes from body language & tone of voice?
  - -55% body language, 38% tone, 7% words popularized by (Mehrabian 1971)



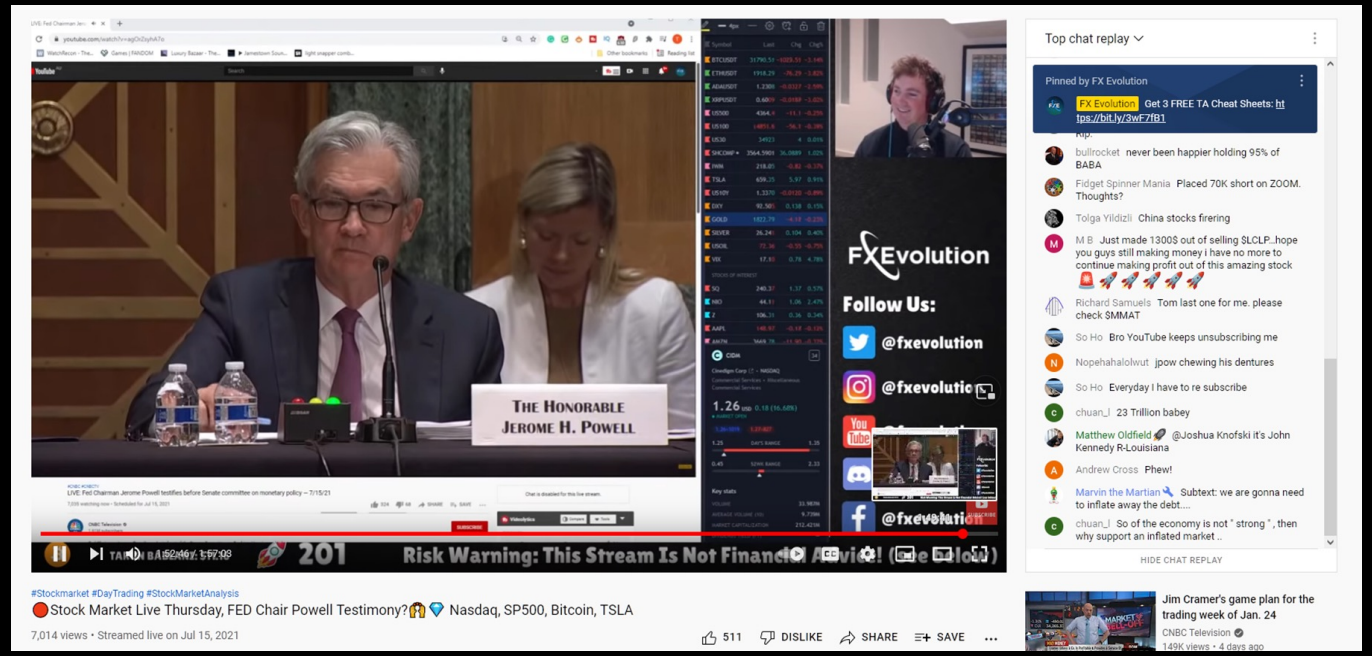
# WE EXPLORE WHAT IMPACTS (IF ANY) FACIAL EXPRESSIONS, TONE OF VOICE AND WORDS CHOSEN HAVE ON MARKETS?



# WHY BODY LANGUAGE & TONE OF VOICE MIGHT MATTER…

1. Traders (institutional or retail) watch communication and either:
   i. Signals directly affected their trades (i.e., actively watching for signals)
   ii. Signals impact traders sentiment/confidence which affects their trading behaviour

**Direct Channel**



# WHY BODY LANGUAGE & TONE OF VOICE MIGHT MATTER…

2. Media reports on the event they watched and the soft forms of communications influenced their message/coverage of events *AND* the behaviour of traders watching is impacted (consciously or unconsciously)

"…So we all heard what Yellen was saying, but what about her body language…in one of the clips you watched, ***her arms cross, the eyebrows go up and that's a sign of insecurity….she was much more secure when talking about monetary policy than financial regulatory policy*** … every time she tried talked, she got cut off, ***you saw the frustration and anxiety***… If we did this with Bernanke, there were times Bernanke's ***voice got so shaky*** …there were."

- CNBC Streets Signs (February 11, 2014).

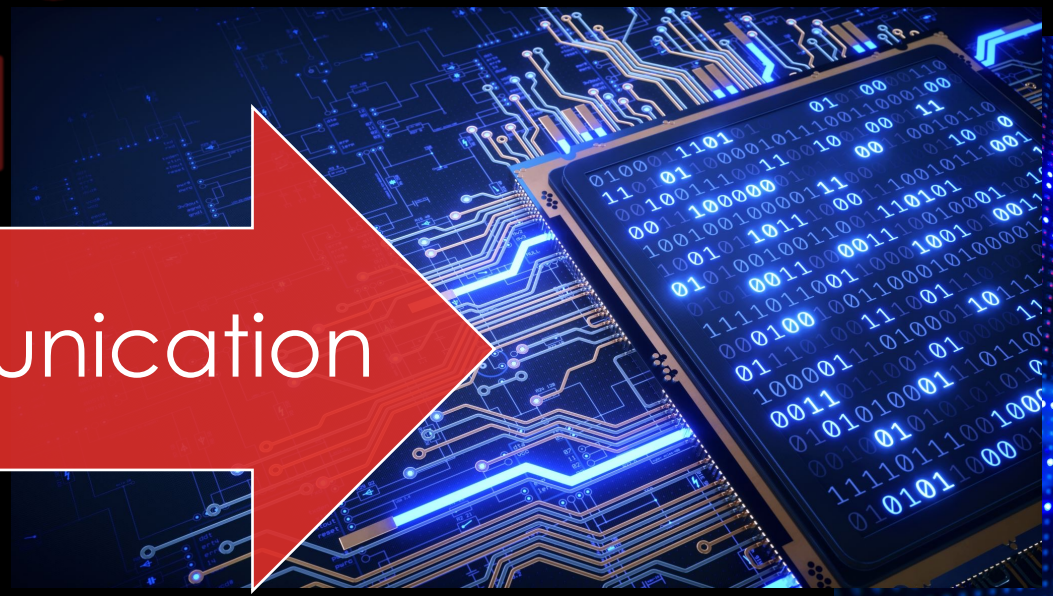
**Indirect Channel**

# WHY BODY LANGUAGE & TONE OF VOICE MIGHT MATTER…

3. Algorithmic trading uses these signals as inputs to algorithms

**Automated Channel**

Fed Communication



# *TO STUDY THE EFFECTS OF THE MESSAGE & THE MESSENGER WE NEEDED ACCESS TO:*

| Traditional data | Non-traditional data (*often in copyright*) | Lots of storage (behind firewalls) | Ample computing power | Tools (and/or RAs to program!) |
|---|---|---|---|---|
| •E.g., stock market data) | •Text<br>•Video | •frame info (~29.97 per sec)<br>•Tick-by-tick stock data | •GPU capability<br>•Lots of memory | •Open source<br>•Software packages |

# THE PROJECTS' DATA SOURCES

## From Internet Archive

- TV Archive Content
  - CSPAN
  - CNBC
  - Bloomberg TV
  - House & Senate Videos

## Other Sources/Providers

- Other Videos:
  - CSPAN Archive + Youtube + Gov't Archives
- Textual
  - CQ RollCall + Lexis-Nexus + HathiTrust + Factiva + Twitter
- Financial Data
  - TickData + Refinitiv + CME + WRDS (TAQ)

# WHY DO WE FOCUS ON TESTIMONIES?

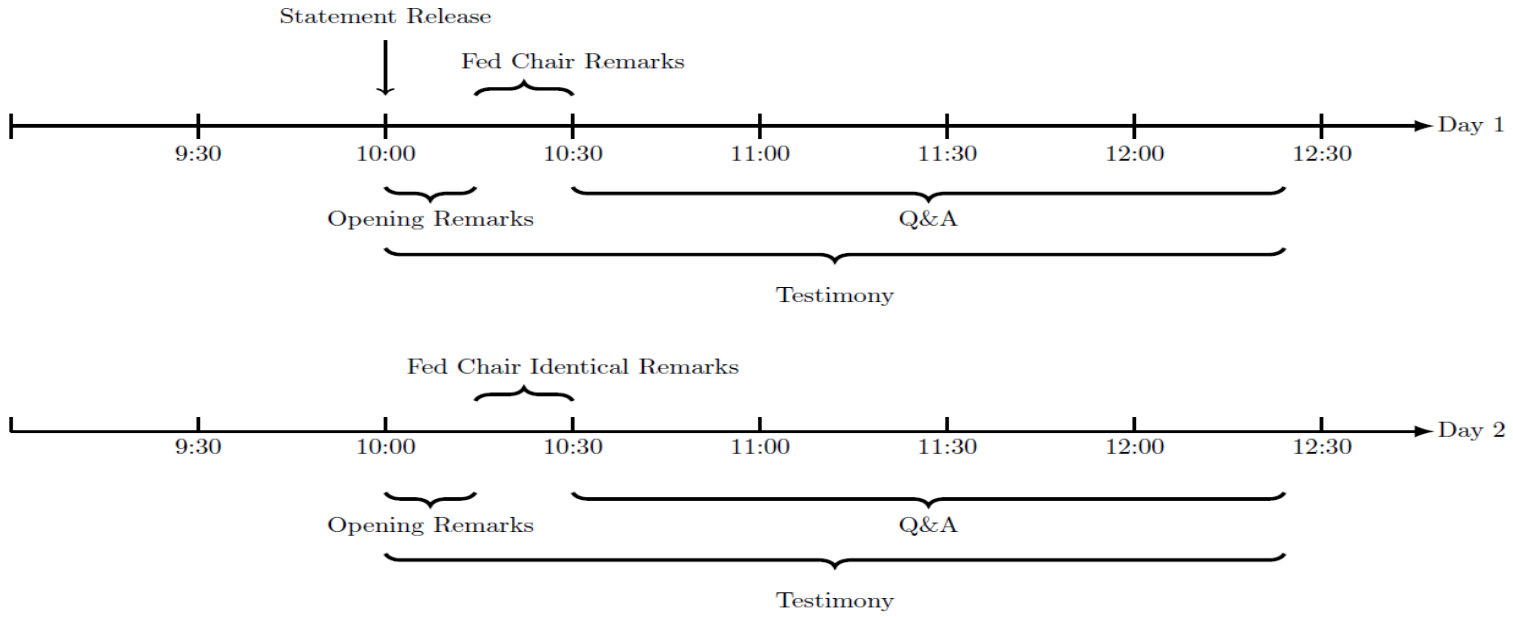
- Don't occur on same day as a policy announcement
- Followed by Fed watchers and investors
- Wide Media Coverage (as large as Press conferences)
  - TV (CSPAN, CNBC, Bloomberg)
  - Newspapers and Wires
  - Bloomberg Terminal

***High quality TV footage & Transcripts are  available in the archives & in our library collections!

# STRUCTURE OF SEMI-ANNUAL TESTIMONIES (~2-3HRS)

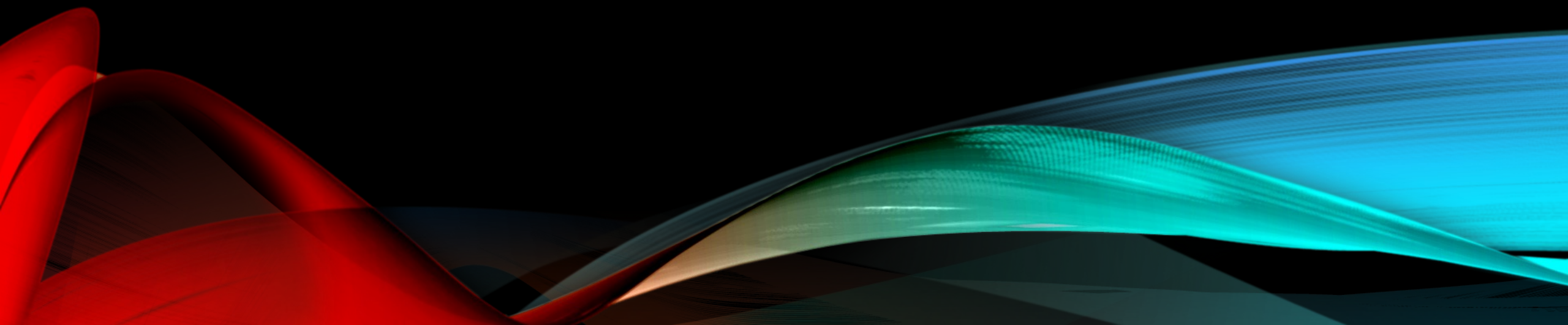
Two congressional testimonies, within a day or two days, alternate

- the House Financial Services committee
- the Senate Banking, Housing, and Urban Affairs committee




**Same remarks given by Fed Chair Day 1 & 2

# HOW DO WE USE THE DATA SOURCE?



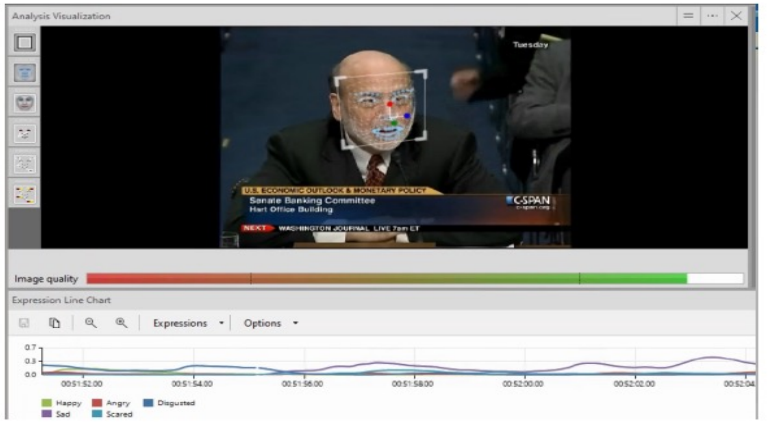
# MEASURING EMOTIONS IN COMMUNICATIONS USING THE RESOURCES & TOOLS

## Testimony data: an example from March 1st, 2011

### From the transcript:

| Speaker | Sentence |
|---|---|
| MENENDEZ: | And so would you give me your view of how the first and second rounds of quantitative easing are working? |
| BERNANKE: | I think they're working – I think they're working well. |
| | The first round in March 2009 was almost – almost the same day as the trough of the stock market. |
| | Since then, the market has virtually doubled. |
| | The economy was going from total collapse at the end of the first quarter of '09 to pretty strong growth in the second half of '09. |
| | And as I said, it's now in the seventh quarter of expansion. |

### From the video:

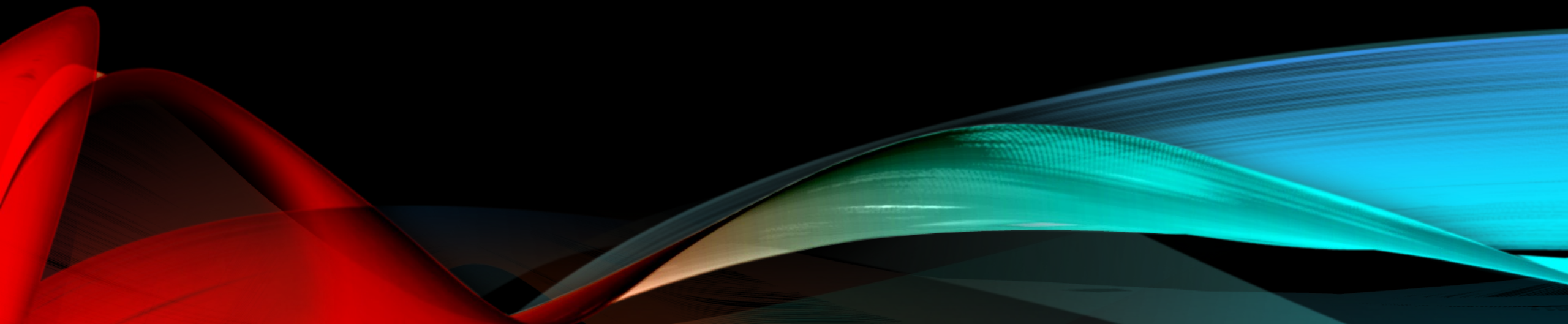


# UNITS FOR ANALYSIS: BLOCKS OF TEXT

- Blocks of text
  - Remark section:
    - 10 sentences (~ 1 min)
    - Typically remark section 15 min
  - Q&A section:
    - by round (~5 min) for full sample
    - By individual question-answer pairs ( >15s)
    - Average testimony 2.5 hours with 21 Q&A blocks

***Block by sentences allows for coherent analysis of content and emotion in statements

# USING THE TEXT



# PRIMARY TEXT-BASED SENTIMENT MEASURE

## BERT Language Model

- Fine-tuned BERT model with labels from 2 testimonies to get sentiment for each sentence

- Fine-tuned model outperformed standard BERT and Fin-Bert (based on F1 Scores)
  - (Sensitivity based on trained version using testimonies outside 32)

# SAMPLE OF AKHZ SENTIMENT LABELS

| | | |
|---|---|---|
| | estimates of that? | |
| Bernanke | It certainly did have some negative effects. | -1 |
| Bernanke | The increased financial concerns led to declines in the stock market, increased credit spreads, and was one of the reasons why we marked down our outlook for the U.S. economy. | -1 |
| Bernanke | That's absolutely right. | 0 |
| Bernanke | I think that, first, I think that situation is improving. | 1 |
| Bernanke | Confidence has been coming back in part because of the Federal Reserve support for the dollar funding markets. | 1 |

# BLOCK AHKZ SENTIMENT MEASURE

- Classified Sentiment for each sentence into Positive (+1), Neutral (0), Negative (-1)

- For testimony t in block b, testimony $\tau$, speaker *i*,

$$TEXT^{i}_{\tau,b} = mean\left(T0^{i}_{\tau,b}\right)/SD^{i}{}_{TEXT}$$

Where $T0^{i}_{\tau,b}$ is the sentiment scores for i's sentences in block b, testimony $\tau$

Text measure is normalized by standard deviation across all testimonies

- Sentiment is calculated for both Fed Chair & Senators/Congressional Representatives

# OTHER TEXT-BASED MEASURES

## Dictionary Based

- Dovish/Hawkish measure
  - Identifies specific keywords in each sentence and uses rules to define hawkish/dovish sentiment
  - Lexicon from Gorodnichenko et al (2022)

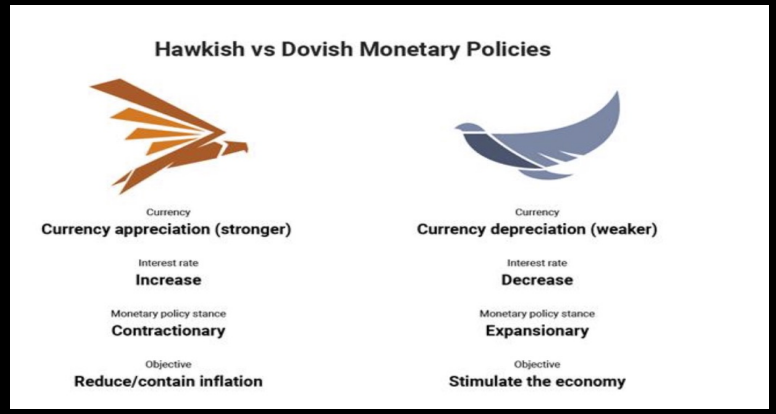
*Table 2. Dictionary for hawkish and dovish words.*

| Panel A1 | Panel A2 |
|---|---|
| inflation expectation, interest rate, bank rate, fund rate, price, economic activity, inflation, employment | anchor, cut, subdue, declin, decrease, reduc, low, drop, fall, fell, decelarat, slow, pause, pausing, stable, non-accelerating, downward, tighten |
| Panel B1 | Panel B2 |
| unemployment, growth, exchange rate, productivity, deficit, demand, job market, monetary policy | ease, easing, rise, rising, increase, expand, improv, strong, upward, raise, high, rapid |
| Panel C | |
| weren't, were not, wasn't, was not, did not, didn't, do not, don't, will not, won't | |

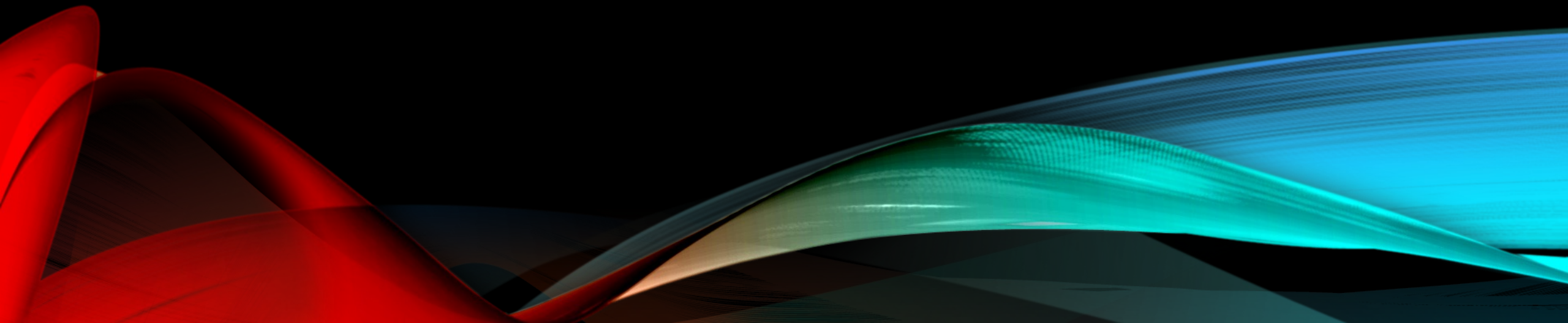
Notes: This table shows the words/phrases used to classify text into dovish/hawkish.

## Bert Fine-Tuned Stance Metric

- Based on combination of labelled data from Gorodnichenko et al (2023) + labelled data from 3 additional testimony transcripts (outside 32)



# VOCAL EMOTION



# CONSTRUCTION OF VOICE MEASURE

Similar approach to Dietrich et al. (2018); Dietrich et al. (2019)

- Sentences timestamps: forced alignment algorithm
- Parse audio to sentences level
- Audio analysis: Praat
- Produce pitch data (F0- fundamental frequency) at 15ms intervals for each person
  - 60 - 180Hz for man, 160 - 300Hz for woman
- High pitch associate with active and intensified emotions

# CONSTRUCTION OF VOICE MEASURE

Block Level Voice-emotion index for speaker i block b, testimony $\tau$, is defined as:

$$VOICE^{i}_{\tau,b} = mean\left(F0^{i}_{\tau,b} - \overline{F0^{i}}\right)/SD^{i}{}_{VOICE}$$

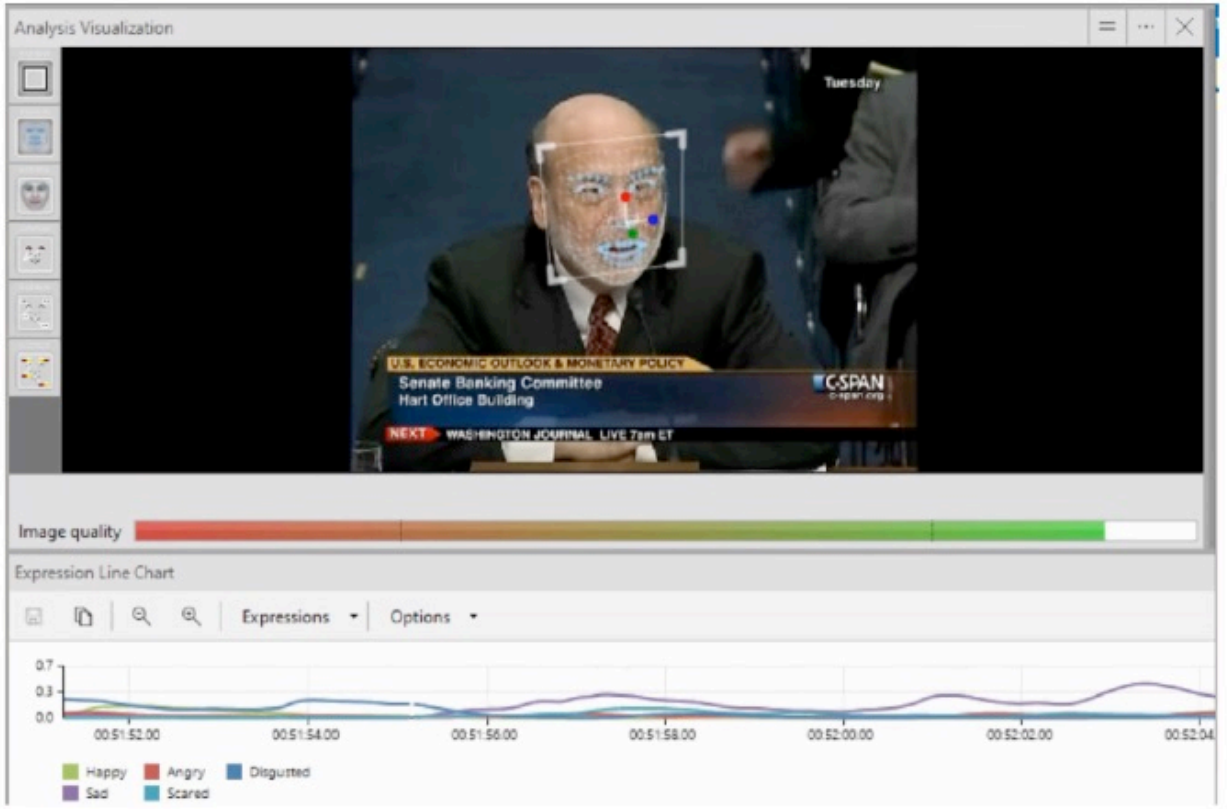
- Measures created for both Fed Chair and Senators/Congressional Representatives
- Demeaned for different groups (Yellen, Bernanke, Senate + Congressional reps)
- Standard deviations for chair and Senate+ Congressional reps calculated across all testimonies

# FACIAL EMOTIONS

# FACEREADER OUTPUT

Emotion data construction (base):

- FaceReader emotions



# OUR PREFERRED APPROACH:

1) Use software to identify facial action units (AUs)

**models to extract facial emotion from Microsoft, Facereader etc. are not trained on people talking

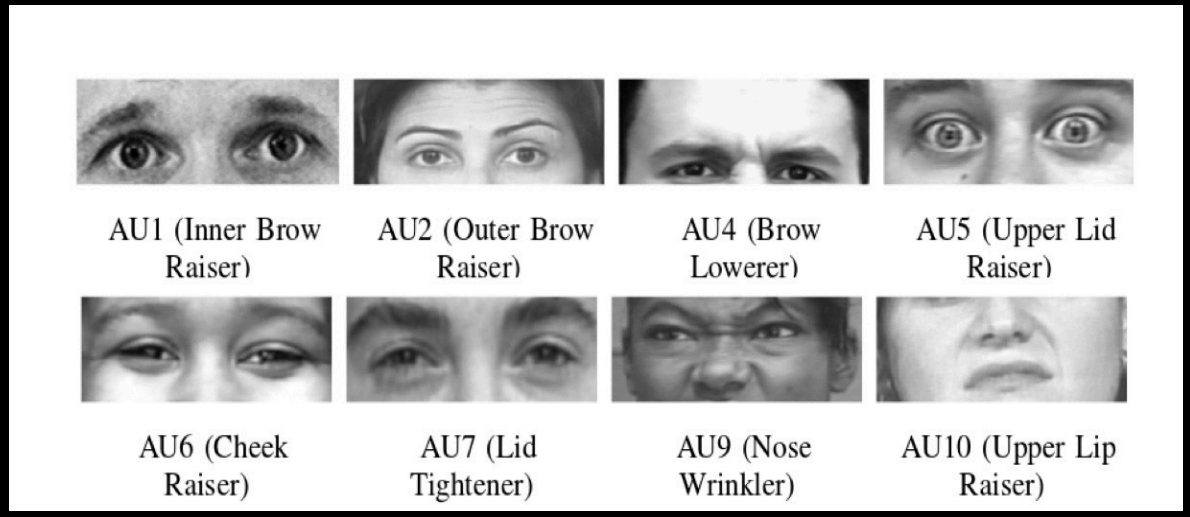
- Manual construct facial emotion data using AUs

# OUR PREFERRED APPROACH:

2) Use mapping from facial AUs to emotions from psychology literature & remove AUs related to the mouth

Facial emotions are the combination of action units

| Emotion | Action Units |
|---|---|
| Sad | 1+4 |
| Fear | 1+2+4+5 |
| Angry | 4+5+7 |
| Disgust | 9 |




3) MS Video indexer (cross-checked manually) to identify 166 participants

# FACIAL EMOTION MEASURE'S CONSTRUCTION

- Frequency: 30 frames per second to capture micro expressions and macroexpressions
- The Face metric is based on the averaged scores of upper FACs associated with negative emotions for i block b, testimony τ, normalized by the standard deviation:

$$FACE^{i}_{\tau,b} = mean\left(Facescore^{i}_{\tau,b}\right)/SD^{i}{}_{FACE}$$

Where $Facescore^{i}_{\tau,b}$ is the set of scores -0.25 (Sad + Angry + Fear + Disgust) for each frame (f) containing the face of individual i in testimony τ

- Measures created for both Chair and Senators/Congressional Representatives
- Standard deviation is calculated across all testimonies

Note: Higher value of index means less negative emotions expressed.

# FINANCIAL MARKET DATA

Tick-by-tick data

Focus on S&P, VIX & Eurodollar futures

# HIGH FREQUENCY ALIGNMENT OF DATA

- Emotion data (3 dimensions) is aligned by sentence
  - Generated timed transcripts at the sentence level from audio and CQ-Rollcall transcripts released the day of testimony (Forced alignment algorithm)
  - Facial metrics merged using timestamps from the corresponding audio

# HIGH FREQUENCY ALIGNMENT OF DATA

- New method for alignment across sources to eliminate noise
- Emotion measures aligned with real time financial data through use of:
  - CNBC live coverage timestamps with S&P real time data onscreen
  - CSPAN live coverage timestamps
  - Testimony official start time & end time

**Result is a dataset that aligns different data very precisely so High Frequency analysis can be done.

# MINE BREAKING NEWS FOR CONTROLS

**Creates a High Frequency news series that is classified by type

We drop the blocks that overlap with market-wide breaking news that are not related to the testimony.

# METHODS: LOCAL PROJECTIONS AND DIFF-IN-DIFF TYPE REGRESSIONS

Local projections

$$Outcome_{t,b+h} - Outcome_{t,b} = \beta^{(h)}_{\text{TEXT}} \text{TEXT}_{tb} + \beta^{(h)}_{\text{VOICE}} \text{VOICE}_{tb} + \beta^{(h)}_{\text{FACE}} \text{FACE}_{tb} + \text{controls} + \text{constant} + \varepsilon^{(h)}_{tb}$$

**Outcome**: the outcome variable (e.g., the log price of S&P 500)
t: testimony date; b: end of block b; b+h: h minutes after the block b
**Controls**:

- testimony fixed effects
- dovish/ hawkish sentiment based on Gorodnichenko et al 2021

**Parameters of interests**: $\beta^{(h)}_{\text{TEXT}}$, $\beta^{(h)}_{\text{VOICE}}$, $\beta^{(h)}_{\text{FACE}}$, $h = 1, 2, 3, \ldots$

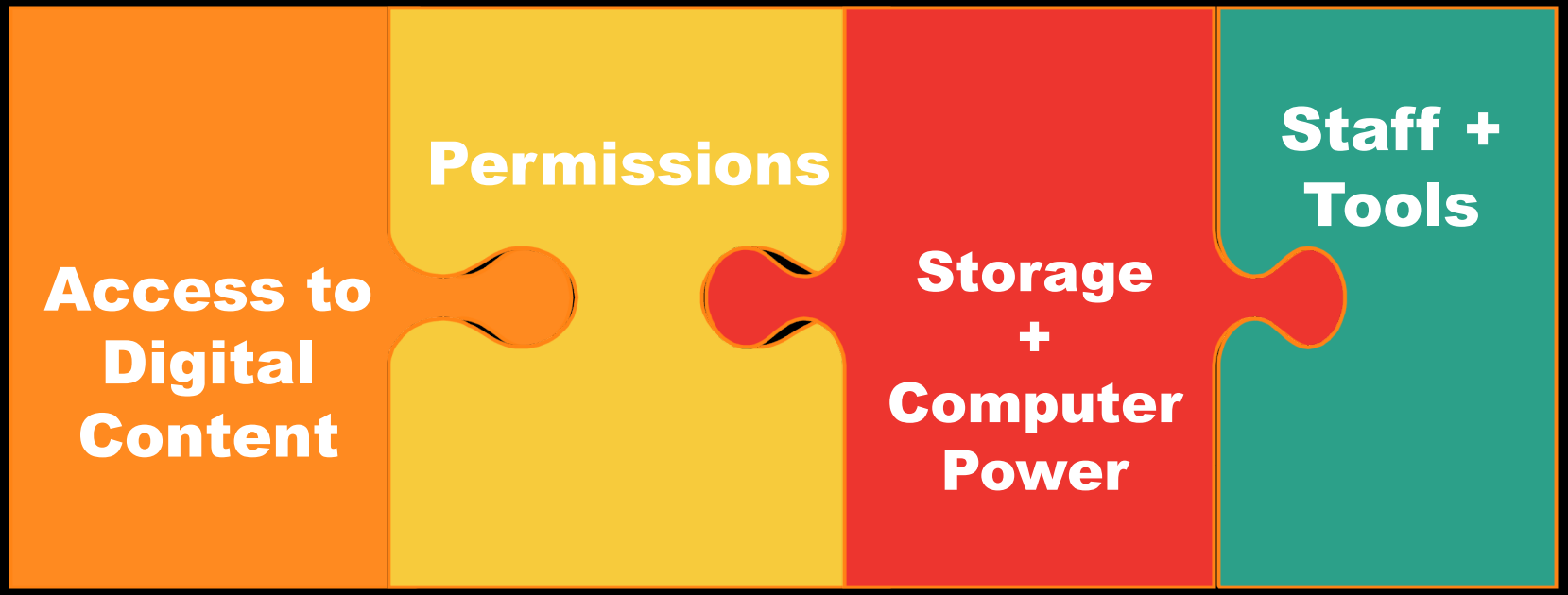
*in Q&A section controls also include fraction Chair's sentences, voice hear and face onscreen and Members text, voice and facial expressions, and lags of the variables.

# SUMMARY OF RESULTS

- Text, voice and facial emotions in testimonies move financial markets (indexes not significantly correlated)
- Magnitude of impacts are different between Remarks and Q&A portions
- Magnitudes of responses grow in the days following the testimony
- Responses to sentiment in text, face and voice can differ significantly across topics
- Responses may also differ by Fed Chair
- Congressional members' emotions can also affect markets
- Using "off the shelf" tools can give unintuitive results

# HOW CAN LIBRARIES SUPPORT THIS TYPE OF RESEARCH?

# Access, Access, Access!



# THANK YOU

## COMMENTS & QUESTIONS WELCOME

Link to working paper: ***More Than Words**: Fed Chairs' Communication During Congressional Testimonies*

LINK: https://ideas.repec.org/p/tor/tecipa/tecipa-737.html

**Email: m.alexopoulos@utoronto.ca**