

# Discussion of “Can ChatGPT Decipher FedSpeak”

Robin L. Lumsdaine  
September 29, 2023

# Disclosures / Disclaimers

---

- Was already familiar with paper / authors
- The views expressed are mine alone and do not necessarily reflect the views of these or other coauthors or any of the institutions with which I am affiliated

# Background

---

- Development of LLMs has transformed textual analysis
- Introduction of ChatGPT has revolutionized the field
  - Lower barrier to entry
  - Significant buzz/attention/consternation

# This paper

---

- Creatively uses ChatGPT to identify and interpret language used in Fed speeches
- Classification task: monetary policy stance
- Compares a variety of models
- Demonstrates use in identifying macro shocks; compares to Romer and Romer (1989, 2023)
- Great job documenting prompts and use of ChatGPT for this task

# Romer & Romer (*AER*, 2023)

---

- “Features of Good Narrative Analysis
  - A reliable narrative source
  - A clear idea of what one is looking for in the source
  - Approach source dispassionately and consistently
  - Document narrative evidence carefully”

# Models compared

---

- Manually labelled benchmark (“the Bryson model”)
- LLMs: GPT-3 and BERT
- Dictionary-based methods: LM, Henry, NRC
- Classification task: Hawkish, Mostly Hawkish, Neutral, Dovish, Mostly Dovish
- Sentence-level analysis; uniform draw of 500 of these

# Questions / comments

---

- Zero-shot and fine-tuning for both GPT-3 and BERT?
- How are “correct” predictions defined?
- Humans have a tendency to say “neutral” when unsure
  - LLM has tendency to overclassify?
  - Incorporate tendency in training?

# Comparison to Romer & Romer

---

- Slightly different corpora, focus only on contractionary shocks
- Are the Romer & Romer papers in the GPT training corpus?
- Robustness with respect to prompt?
- Order of document processing?



# Summary and future ideas

---

- Automate interpretation, leading to reduction in investor disagreement / more accurate predictions
  - Is reduction in investor disagreement desirable?
  - Clarity of message but still heterogeneity of reaction to message
  - Does better understanding  $\Leftrightarrow$  better prediction?
- How does stance relate to vote/decision?
- Time dimension
- Great paper – lots of interesting applications / possibilities