

Do FOMC Minutes Signal Future Rate Moves?

AI vs. Lexicon Sentiment Evidence, 1993-2025

Ian Phiri Chigada, M.A. in Mathematics

FACULTY MENTORS: Yuri Katz, Ph.D.



Katz
Katz School
of Science and Health

1. Research Question

- Modern monetary policy works through both rate decisions and communication.
- This study tests whether the tone of FOMC meeting minutes contains measurable information about future changes in the federal funds target rate.
- Main question: does AI-based sentiment improve on a transparent lexicon-based benchmark?



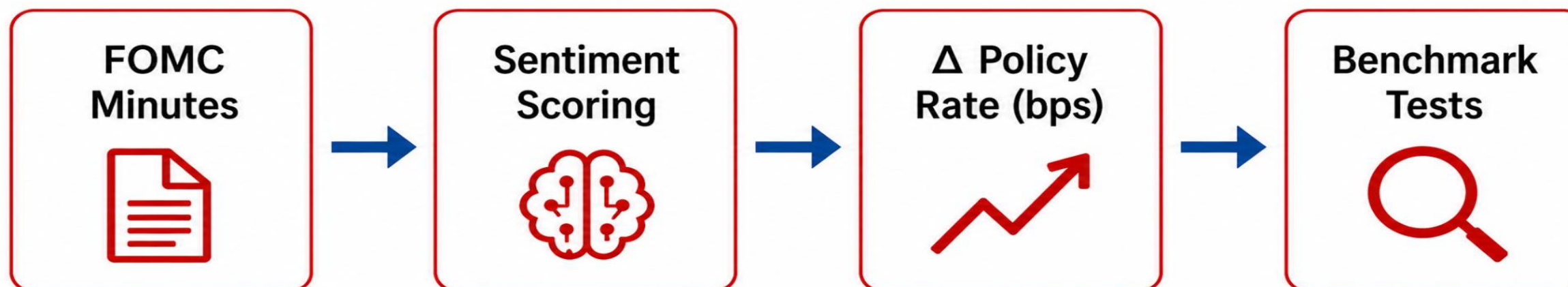
259 FOMC meetings

1993–2025

79 meetings with non-zero $\Delta(\text{bps})$

2. Data & Methods

- Text source: FOMC meeting minutes for scheduled and major unscheduled meetings.
- Rate data: FRED and Federal Reserve historical records.
- AI Sentiment: OpenAI ChatGPT-4o, fixed prompt, 1 = very dovish to 10 = very hawkish.
- Benchmark: Lexicon Sentiment using dictionary-based word counts.
- Tests: correlations, OLS, positive-lag CCFs, distributed-lag regressions, and non-QE robustness checks.



References

- [1] Board of Governors of the Federal Reserve System. Minutes of the FOMC.
- [2] Federal Reserve Bank of St. Louis. FRED Database.
- [3] Newey, W.K. & West, K.D. (1987). Econometrica.
- [4] Hansen, S. & McMahon, M. (2016). Journal of International Economics.
- [5] Loughran, T. & McDonald, B. (2011, 2024).

3. Benchmark Design

Benchmark	Purpose	Key Value
Lexicon Sentiment	Transparent dictionary benchmark	Corr. with $\Delta(\text{bps}) = 0.542$
AI Sentiment	Context-aware LLM benchmark	Corr. with $\Delta(\text{bps}) = 0.556$
Lexicon (non-QE)	Robustness outside ZLB/QE regime	Corr. with $\Delta(\text{bps}) = 0.570$
AI (non-QE)	Robustness outside ZLB/QE regime	Corr. with $\Delta(\text{bps}) = 0.590$

AI outperforms the lexicon benchmark modestly, not dramatically.

Robustness Check (Non-QE Sample)

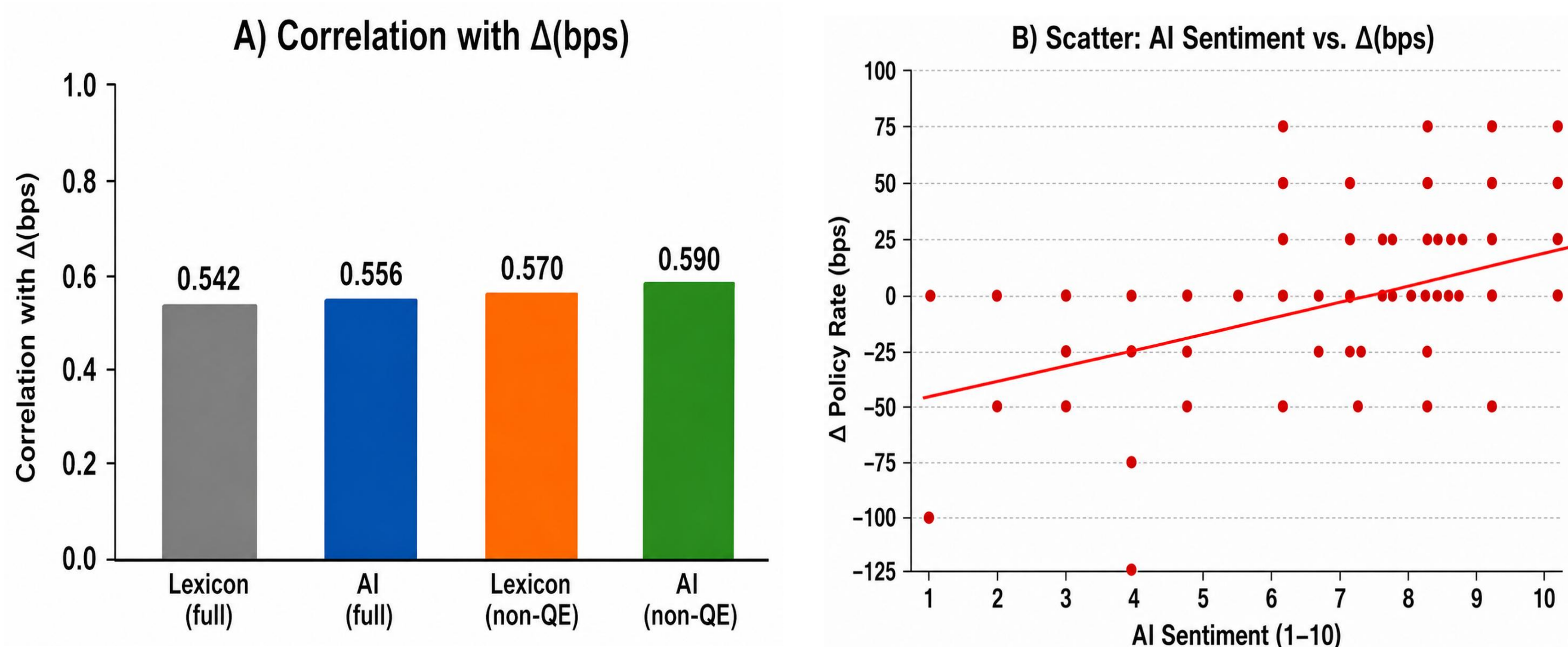
Outside the 2009–2015 QE / zero-lower-bound period, correlations strengthen

(Lexicon: 0.542 → 0.570; AI: 0.556 → 0.590).

Conclusion & Next Steps

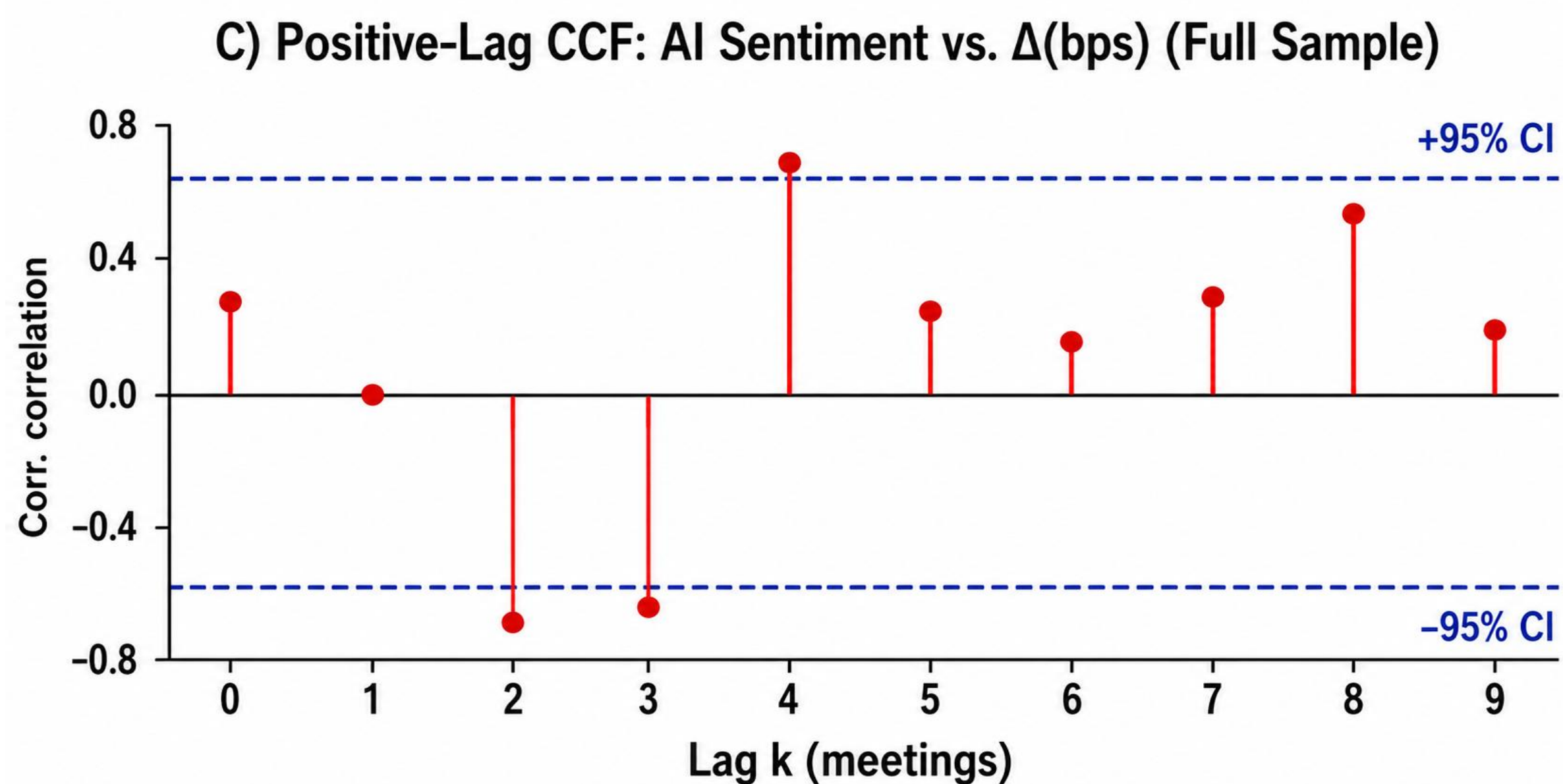
FOMC minutes contain measurable policy information. AI sentiment slightly outperforms lexicon sentiment, especially outside QE, but the improvement is modest and should be treated as a benchmarked signal rather than a stand-alone forecasting model. Future work should compare multiple named LLMs and test out-of-sample performance.

4. Key Results



AI sentiment is slightly stronger than the lexicon benchmark, especially outside the QE period.

More hawkish minutes are associated with a greater likelihood of rate hikes.



Positive-lag CCF peaks around $k = 4$ and $k = 8$ suggest hawkish tone may precede tightening several meetings later.

- Both sentiment measures are positively associated with $\Delta(\text{bps})$, but AI is consistently slightly stronger.
- The signal is clearer when the 2009–2015 QE / zero-lower-bound period is excluded.
- Pre-whitening weakens the lead-lag pattern, so the predictive signal should be viewed as modest and partly persistence-driven.

Acknowledgements

Thank you to Dr. Yuri Katz and the Katz School of Science and Health for their guidance and support.