LLM FOMC SIMULATION

Developing Multi-Agent Frameworks for Federal Interest Rate Prediction

By Adler Viton, Tianji Rao, Xiyue Zhang, Yuhan Hou and Jeremy Tan

Table of Contents

Overview: Introduction to Federal Funds Rate, why do we care?

Baseline Quantitative Modeling

- Quantitative macro data
- Linear Regression Results

LLM Based Modeling

- Motivations
- Unstructured Input Data

Multi Agent System (FOMC Simulation)

- Basic idea
- Clustering
- Base Model
- Simulations
- Chain of Draft

Backtesting & Results

- Methodology
- Results

Conclusion


Background

Intro to Fed Funds Rate

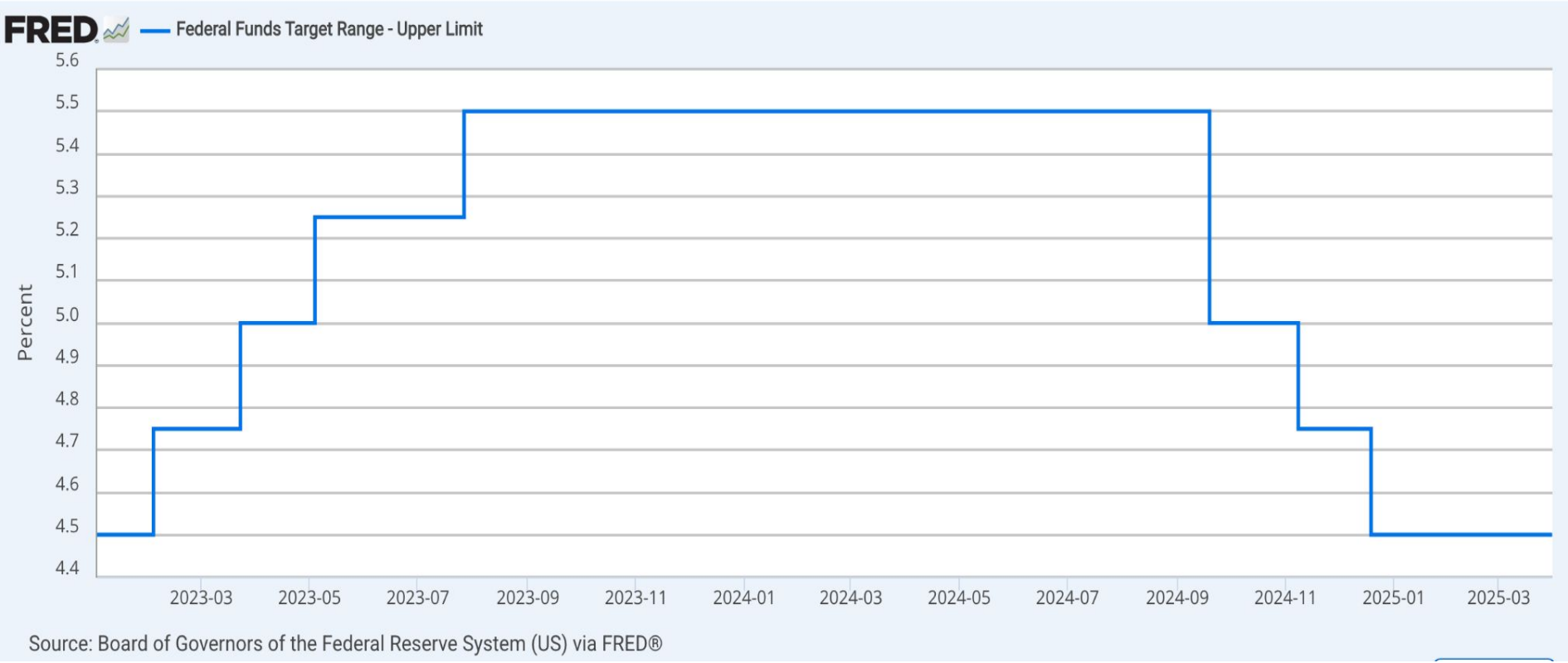
What is the Federal Funds Rate?

- Interest rate at which banks and credit unions lend reserve balances to other depository institutions overnight

Why is it important to BNY?

- Net Interest Margin
- Interest Rate Risk
- Liquidity & Capital Management
- FX Market Conditions

Intro to Fed Funds Rate



Federal Open Market Committee (FOMC) Announcements

Frequency: 8 times a year, plus emergency meetings

Key outcomes:

- **Rate decision**: Announcement of changes to the federal funds rate.
- **Economic assessment**: Insights on employment, inflation, and growth.
- **Forward guidance**: Indications of future rate moves or policy shifts.

What is the Problem?

Financial markets react strongly to decisions made during Federal Open Market Committee (FOMC) meetings, yet **accurately predicting** their impact remains a challenge.

How can we both predict and interpret the Federal Funds Rate to support BNY's risk management strategies?

Who Votes in the FOMC?

The 7 Federal Reserve Governors Board Members

New York Fed President

4 Other Regional Fed Reserve Bank Presidents

THE FEDERAL RESERVE BOARD OF GOVERNORS

Jerome Powell
Chair

Philip Jefferson
Vice Reserve Chair

Michael Barr
Vice Supervision Chair

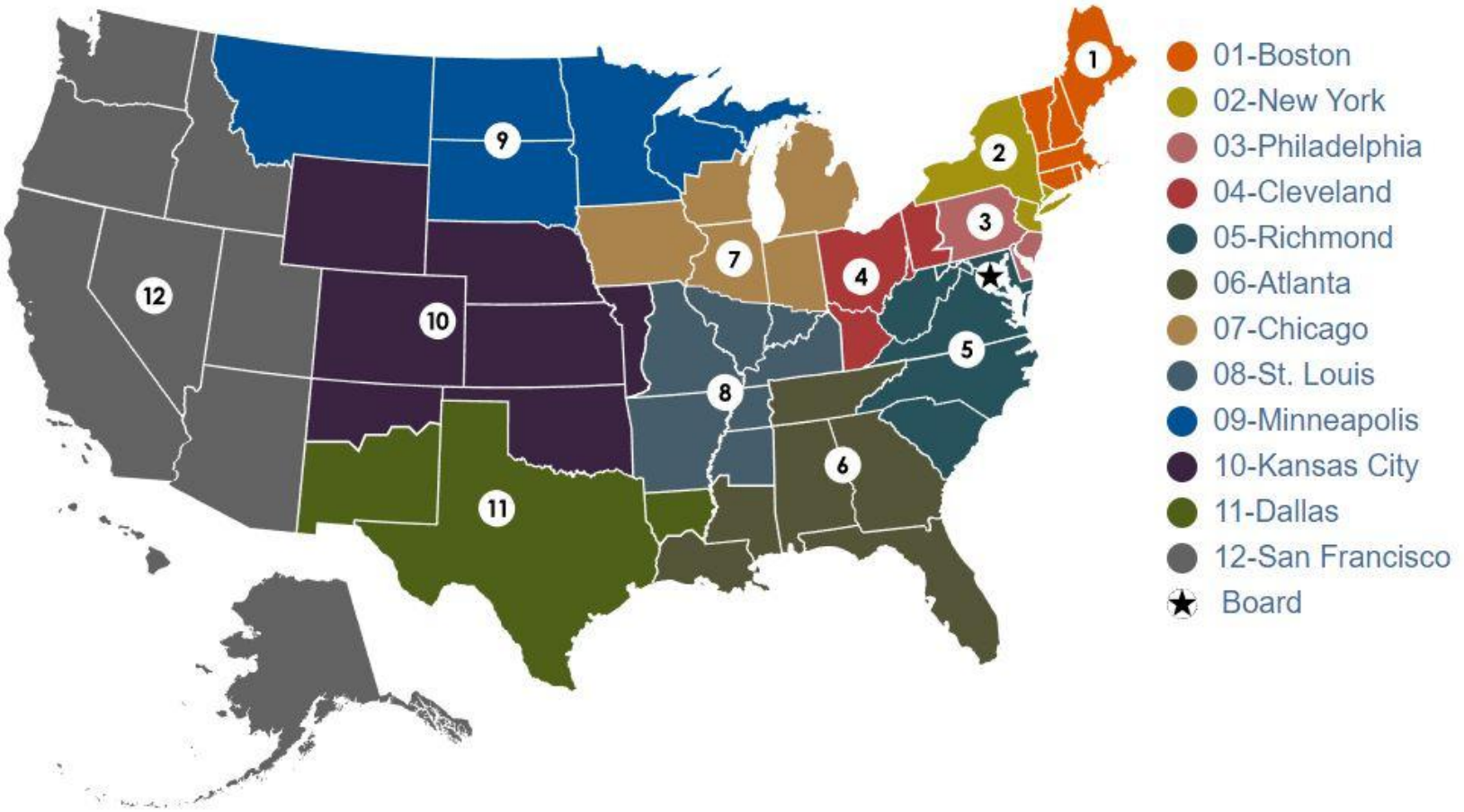
Lisa Cook
Democratic

Christopher Waller
Republican

Adriana Kugler
Democratic

Michelle Bowman
Republican

12 Federal Reserve Banks



Baseline Model

Quantitative Predictors

Predictors for Interest Rate

Inflation

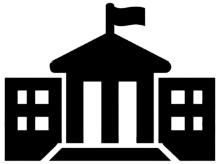
- Personal consumption expenditures price index (PCE)
- Consumer Price Index (CPI)
- MICH: Inflation Expectation

Monetary

- TB3Ms: 3-Month Treasury Bill
- TB6Ms: 6-Month Treasury Bill
- M2: Money Supply Metric

Economic Activity and Growth

- Brave-Butters-Kelley Real Gross Domestic Product
- Volatility Index (VIX)
- Unemployment Rate



Political Indicator Variables

- Who is Fed Chair?
- Which party controls the White House?

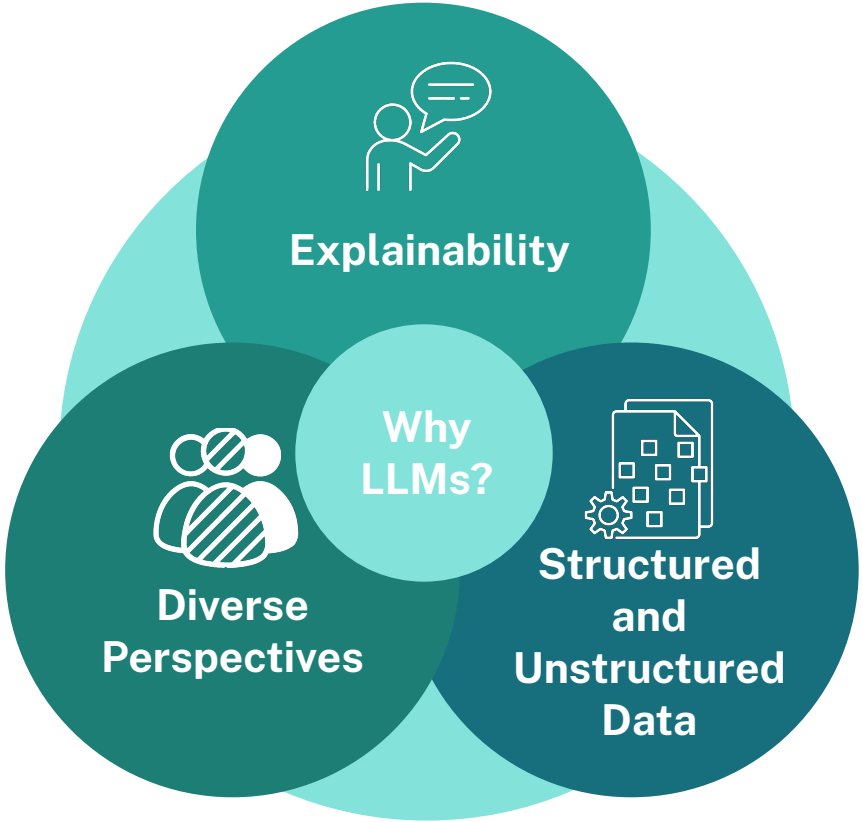
Past Rates

- Current Rate value Before Meeting
- Value of rate change at prior meeting

Baseline Model: Linear Regression

- **Objective**: Predict **rate changes** (e.g., +0.25%, -0.50%)
- **Testing Period**: **July 2022 – March 2025**
- **Prediction Format**: Rounded to nearest **0.25%**
- **Model Accuracy**:
  - **57.14%** on correctly predicting rate changes
  - **33.3%** on “Balanced Test Set”

Why LLMs?



We want to focus on the **decision-making process**, not just the **outcome**

Input Data

Initial Predictors

Predictors for Interest Rate:

Inflation

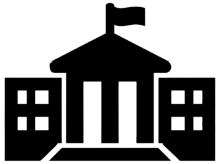
- Personal consumption expenditures price index (PCE)
- Consumer Price Index (CPI)
- MICH: Inflation Expectation

Monetary

- TB3Ms: 3-Month Treasury Bill
- TB6Ms: 6-Month Treasury Bill
- M2: Money Supply Metric

Economic Activity and Growth

- Brave-Butters-Kelley Real Gross Domestic Product
- Volatility Index (VIX)
- Unemployment Rate



Political

- Who is Fed Chair?
- Which party controls the White House?

Past Rates

- Current Rate value Before Meeting
- Value of rate change at prior meeting

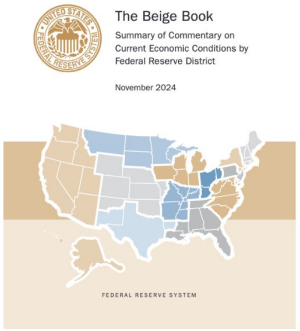
New Data: Beige Book, Dot Plots, Probs

What is the Beige Book?

About 50 pages, released 2 weeks before each FOMC Meeting. Includes a **national** summary, and then one section for each of the **12 federal reserve districts**

***Qualitative* summary of economic conditions:**

Labor Markets
Prices
Consumer Spending
Manufacturing
Real Estate
Community Conditions/Perspectives
Other



Dot Plots

Individual Federal Reserve Committee member's projections for the federal funds rate over the next few years and the longer run

Rate Cut/Hike Probabilities

The probability estimated by represents the market's expectation of a rate hike or rate cut at the next FOMC meeting, derived from the pricing of 30-Day Federal Funds futures

Multi-Agent System Concept

Basic Functionality

Where did this idea come from?

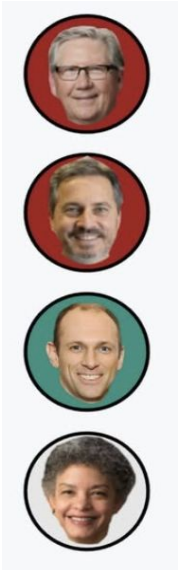
- This concept has been used in simulating judicial rulings from the Supreme Court.
- Interpersonal interactions shape the FOMC's decision making process

Define the Committee of Agents	Give Pertinent Information	Facilitate Agent's Discussion	Final Rate Voting Process
Define a committee of agents that represents the members of the FOMC	Give the Agents the necessary input data, representative of what the FOMC considers	Mimic the discussion dynamics of humans by analyzing and learning from each other	Each Agent eventually takes a final vote on what the Fed's path forward should be

Clustering FOMC Members

Why Cluster Agents:

- **Downsize the number of agents to reduce time, lower costs, and improve discussion readability**
- **Clustering algorithms provide a structured way to group similar agents together**



...

KNN

Features:

1. **Political Party**
2. **Gender**
3. **Tenure Years in their current positions**

...

Regional Pragmatists

Composed of keyboard figures who work closely with local economies

Academic Balancers

Consists of members with strong analytical and academic backgrounds

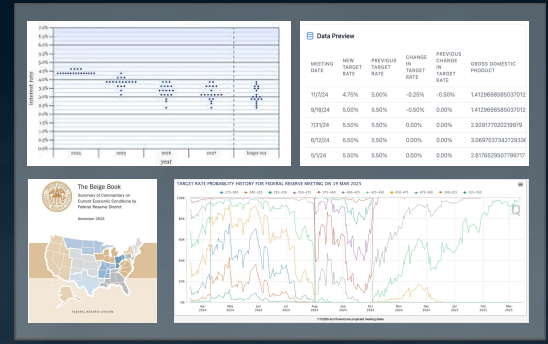
Central Policymaker

Made up of senior decision makers within Federal Reserve system

Model 1: Base Model

Model 1: Sequential Architecture “Base”

Input Data

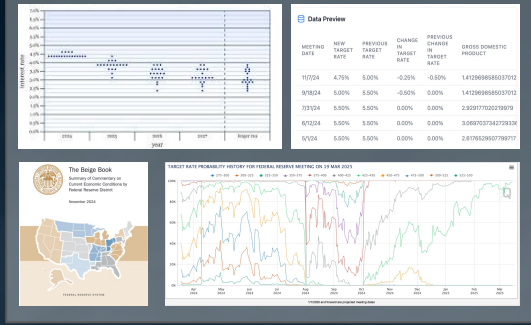


Model 1: Sequential Architecture “Base”

Analyst

Analysis of Implied Probabilities

Input Data



Model 1: Sequential Architecture “Base”

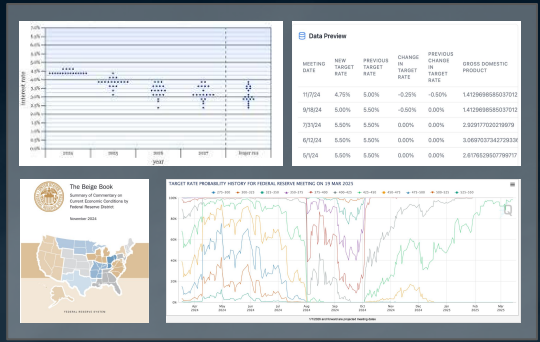
Economist

Generate 3 Potential Solutions

Analyst

Analysis of Implied Probabilities

Input Data



Model 1: Sequential Architecture “Base”

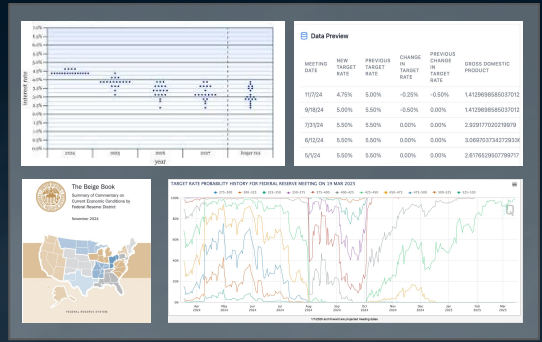
Economist

Generate 3 Potential Solutions

Analyst

Analysis of Implied Probabilities

Input Data



Clustered Member Agents

Analysis

Discussion

Voting

Model 1: Sequential Architecture “Base”

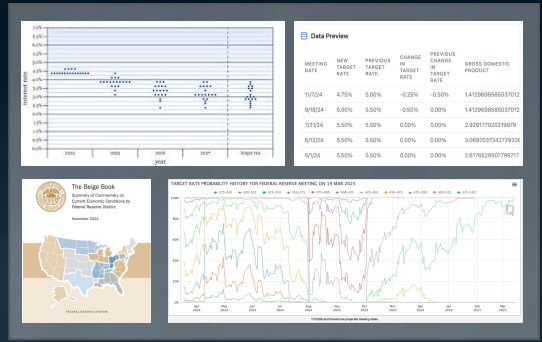
Economist

Generate 3 Potential Solutions

Analyst

Analysis of Implied Probabilities

Input Data



Clustered Member Agents

Analysis

Discussion

Voting

Summary Analyst

Model 1: Sequential Architecture “Base”

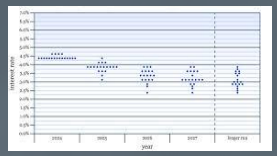
Economist

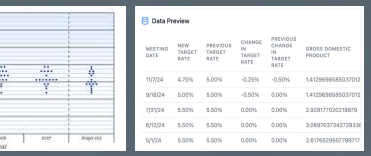
Generate 3 Potential Solutions

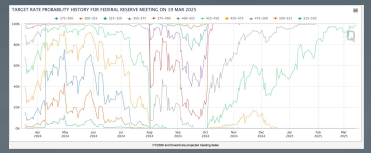
Analyst

Analysis of Implied Probabilities

Input Data







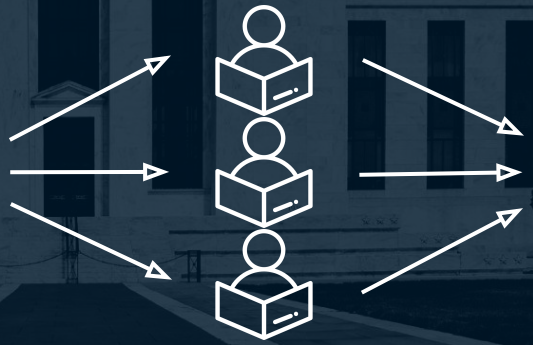
Clustered Member Agents

Analysis

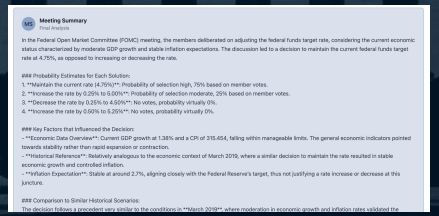
Discussion

Voting

Summary Analyst



Results




Model 2: Base + Simulation

Model 2: Sequential Architecture + Simulation

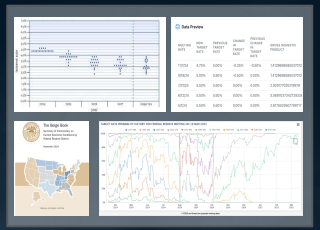
Economist

Generate 3 Potential Solutions

Analyst

Analysis of Implied Probabilities

Input Data



Simulation Memory

Clustered Member Agents

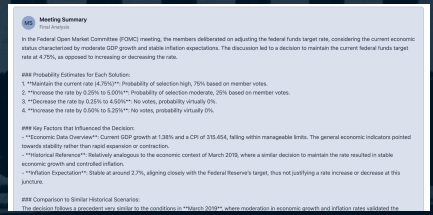
Analysis

Discussion

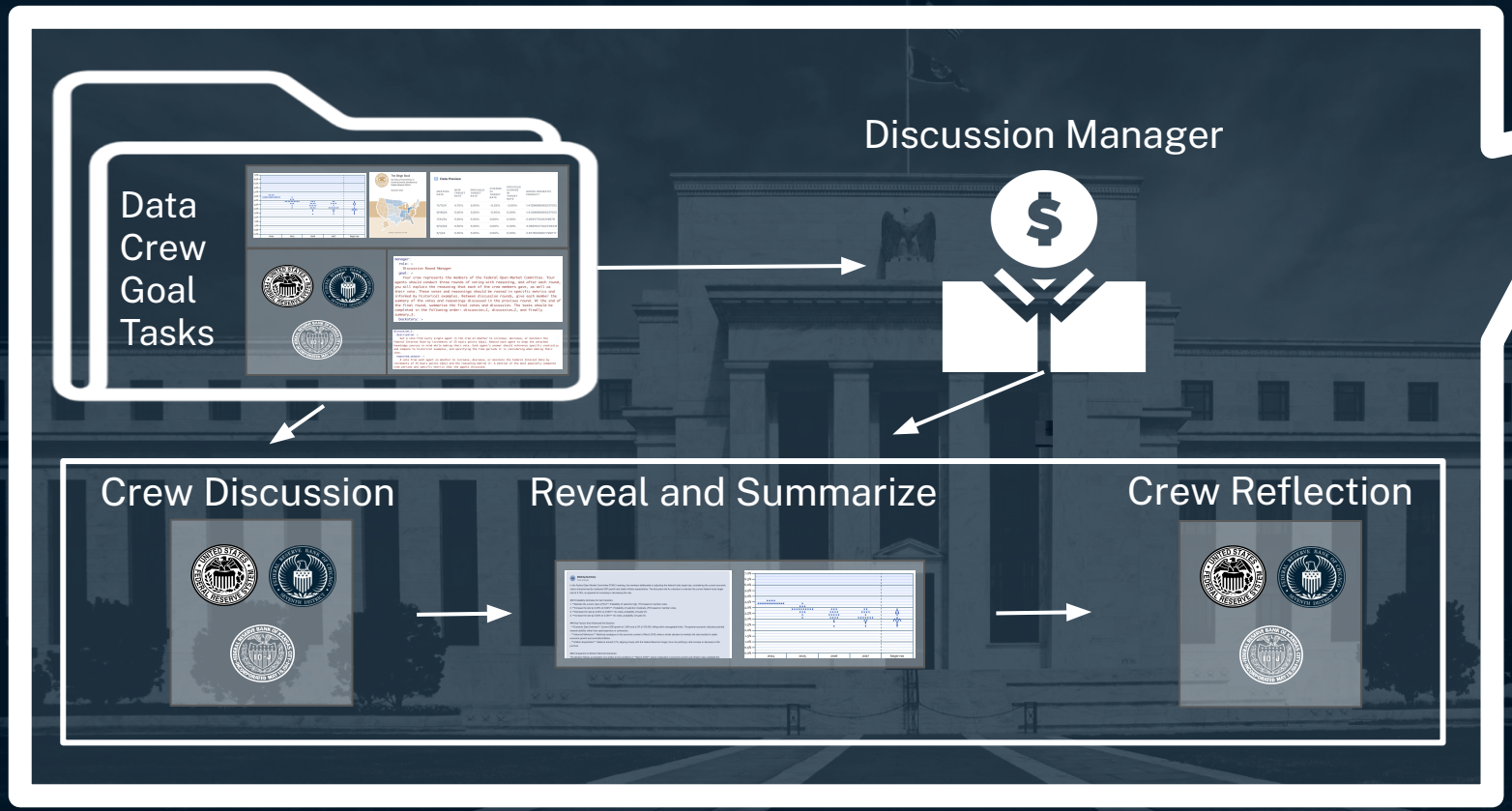
Voting

Summary Analyst

Results



How Do We Get Simulation Memories?



Simulation

Meeting Dates: **Jan, 2022** and **Feb, 2023**

In-Context Learning Objectives:

- Identify gaps in reasoning or information use
- Adjust their decision-making heuristics based on feedback
- Improve performance in similar future scenarios

Long-Term Memory Storage Includes:

- Agent-specific discussion histories
- Key macro observations and referenced examples
- Voting decisions and reflections from simulations

Model 3: Chain of Draft

Reducing Token Size

Chain of Thought (COT)

```
Q: Jason had 20 lollipops.  He gave Denny some lollipops.  Now Jason has 12 lollipops.  How many lollipops did Jason give to Denny?
A: Let's think through this step by step:
1. Initially, Jason had 20 lollipops.
2.  After giving some to Denny, Jason now has 12 lollipops.
3.  To find out how many lollipops Jason gave to Denny, we need to calculate the difference between the initial number of lollipops and the remaining number.
4.  We can set up a simple subtraction problem: Initial number of lollipops - Remaining number of lollipops = Lollipops given to Denny
5. Putting in the numbers: 20 - 12 = Lollipops given to Denny
6. Solving the subtraction: 20 - 12 = 8
Therefore, Jason gave 8 lollipops to Denny.
####
8 lollipops
```

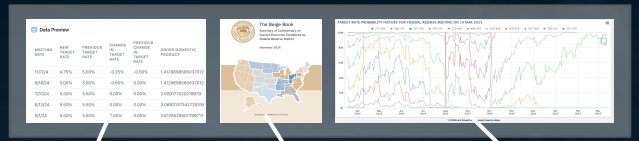
Chain of Draft (COD) Prompting

```
Q: Jason had 20 lollipops.  He gave Denny some lollipops.  Now Jason has 12 lollipops.  How many lollipops did Jason give to Denny?
A: 20 - x = 12; x = 20 - 12 = 8. #### 8
```

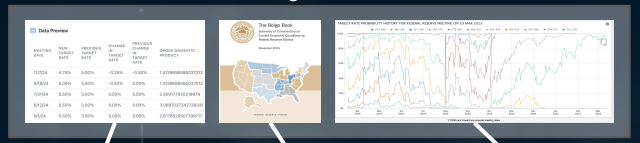
1. Figure adapted from Xu et al. (2025), arXiv:2502.18600

Model 3: Sequential CoD Architecture

Economist analysis of each source



Member's analysis of each source



Each source is looked at in isolation by both economist and each member before options and discussion

Economist Generates Options

Clustered Member Agents

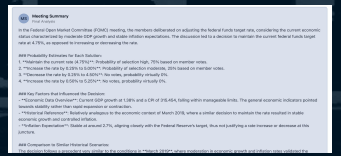
Consolidated Analysis

3-Member discussion

Final Voting Round

Summary Analyst

Results



Backtesting

Evaluation Metrics

For each meeting:

Rate Change Accuracy
Does the model predict the right rate change decision on each meeting?

Individual Rate Change Accuracy
Does each agent vote correctly on each meeting?

Voting Stability
How stable are votes from each agent across simulation runs?

For overall performance:

Whether the model correctly predicts the overall rate decision for each meeting.

$$\text{Average Rate Change Accuracy} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\left(\widehat{\text{Vote}}_i = \text{Vote}_i^*\right)$$

Whether each agent correctly predicts the true rate decision in each meeting.

$$\text{Average Individual Voting Accuracy} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{K}\sum_{k=1}^{K} \mathbb{1}\left(\widehat{\text{Vote}}_{i,k} = \text{Vote}_i^*\right)\right)$$

Evaluate the consistency of model predictions for individual agent votes across simulation runs for each meeting.

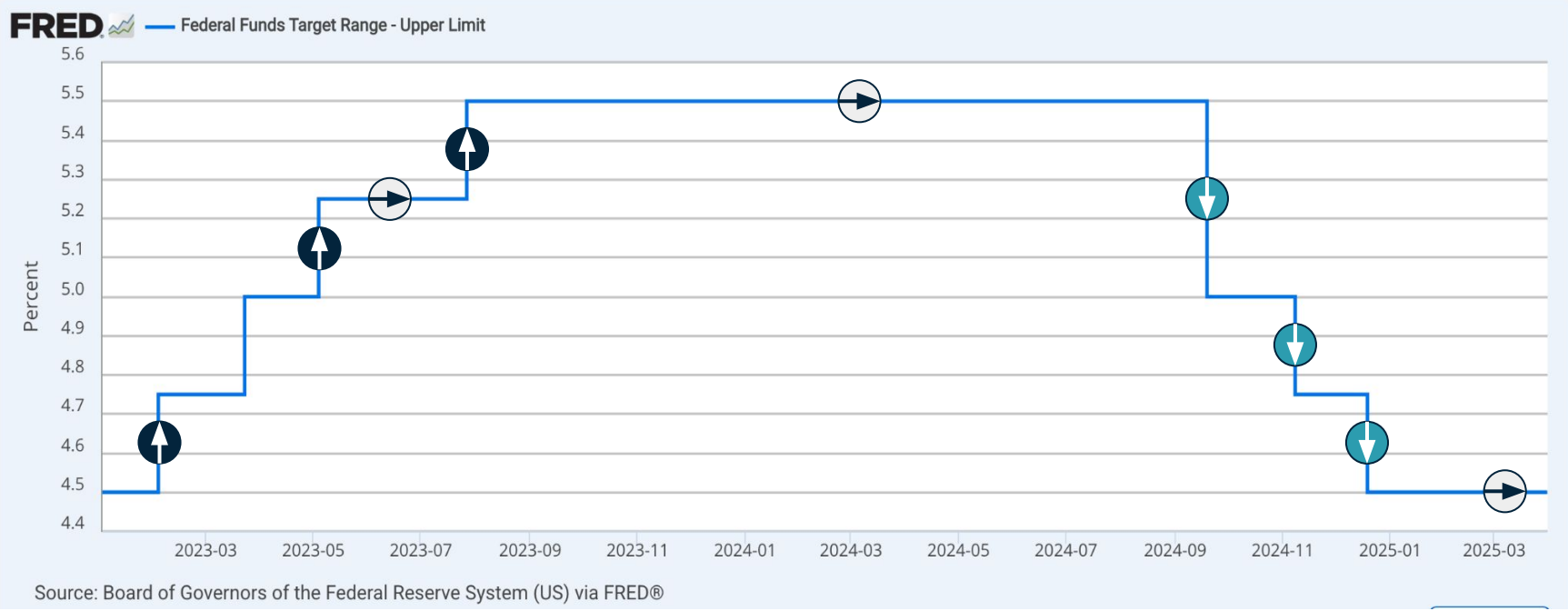
$$\text{Average Voting Stability} = \text{avg}_{i,j,k}\ \mathbb{1}\left(\widehat{\text{Vote}}_{i,j,k} = \widehat{\text{Vote}}_{i,k}^{\text{mode}}\right)$$

Test Meetings

9 total meetings from 2023 and 2024

3 hikes, 3 cuts, and 3 decisions to maintain

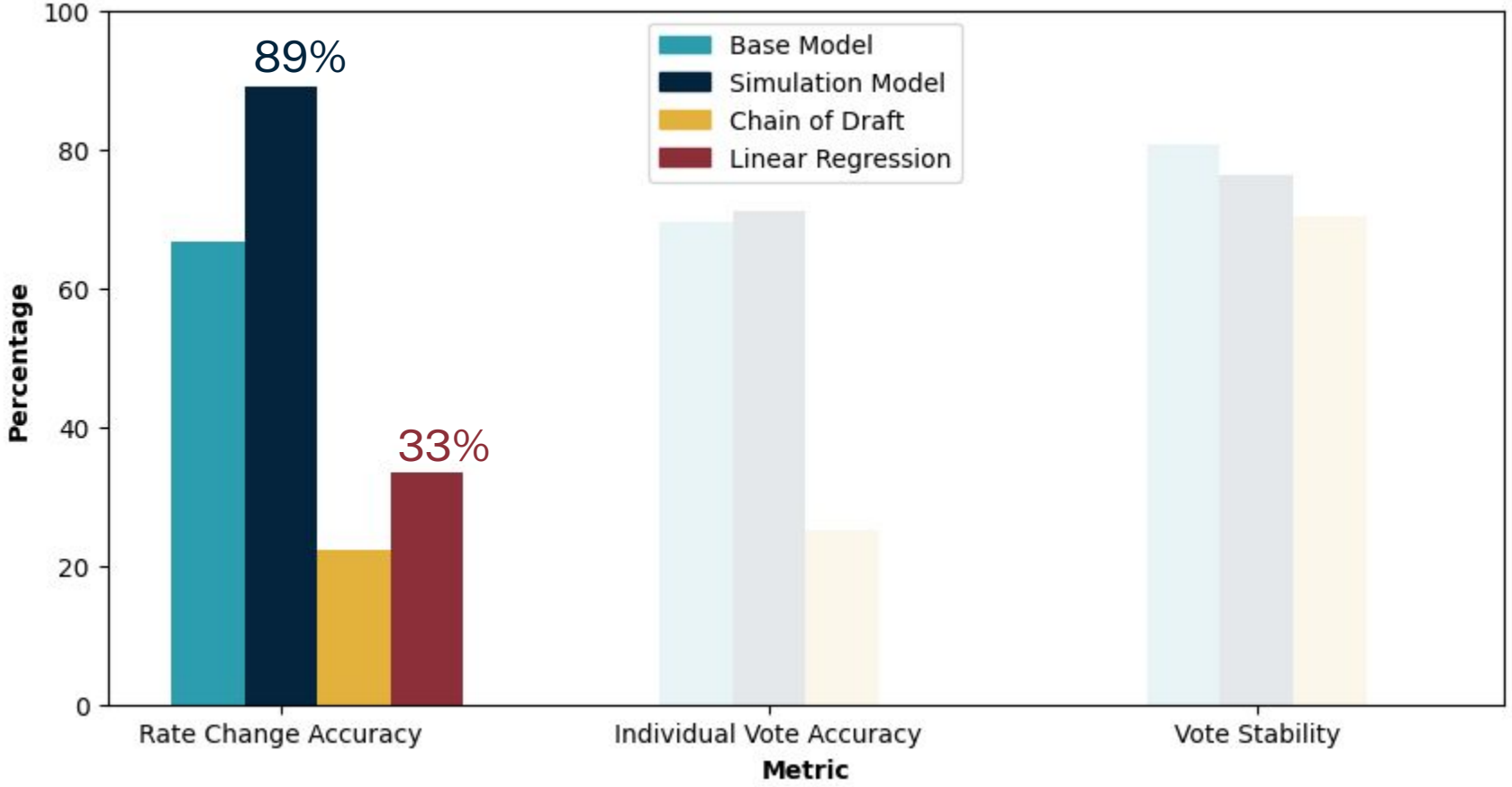
5 runs on each meeting




Results

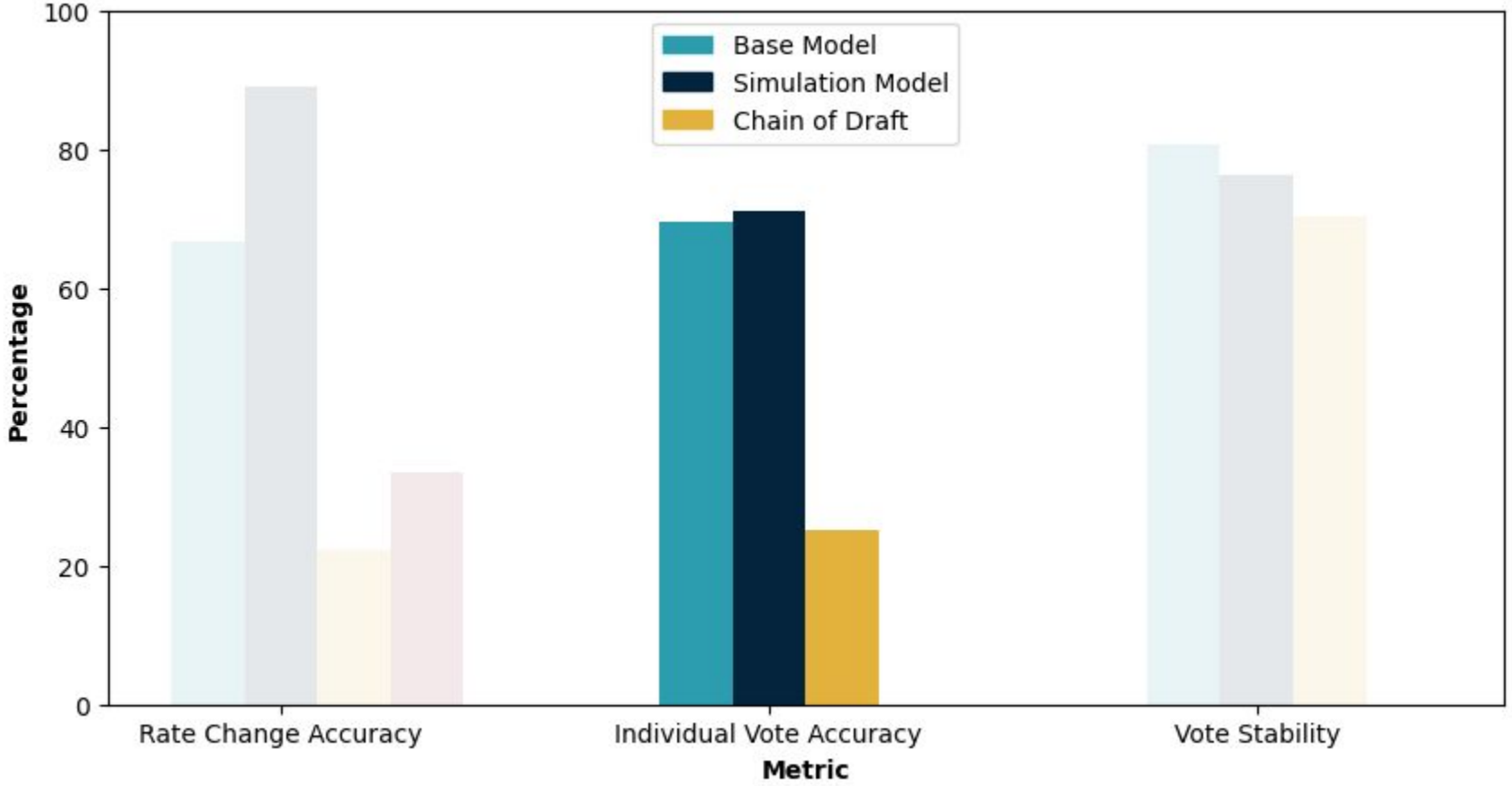
Results

Simulation Model Improves on Linear Regression Accuracy by 56%



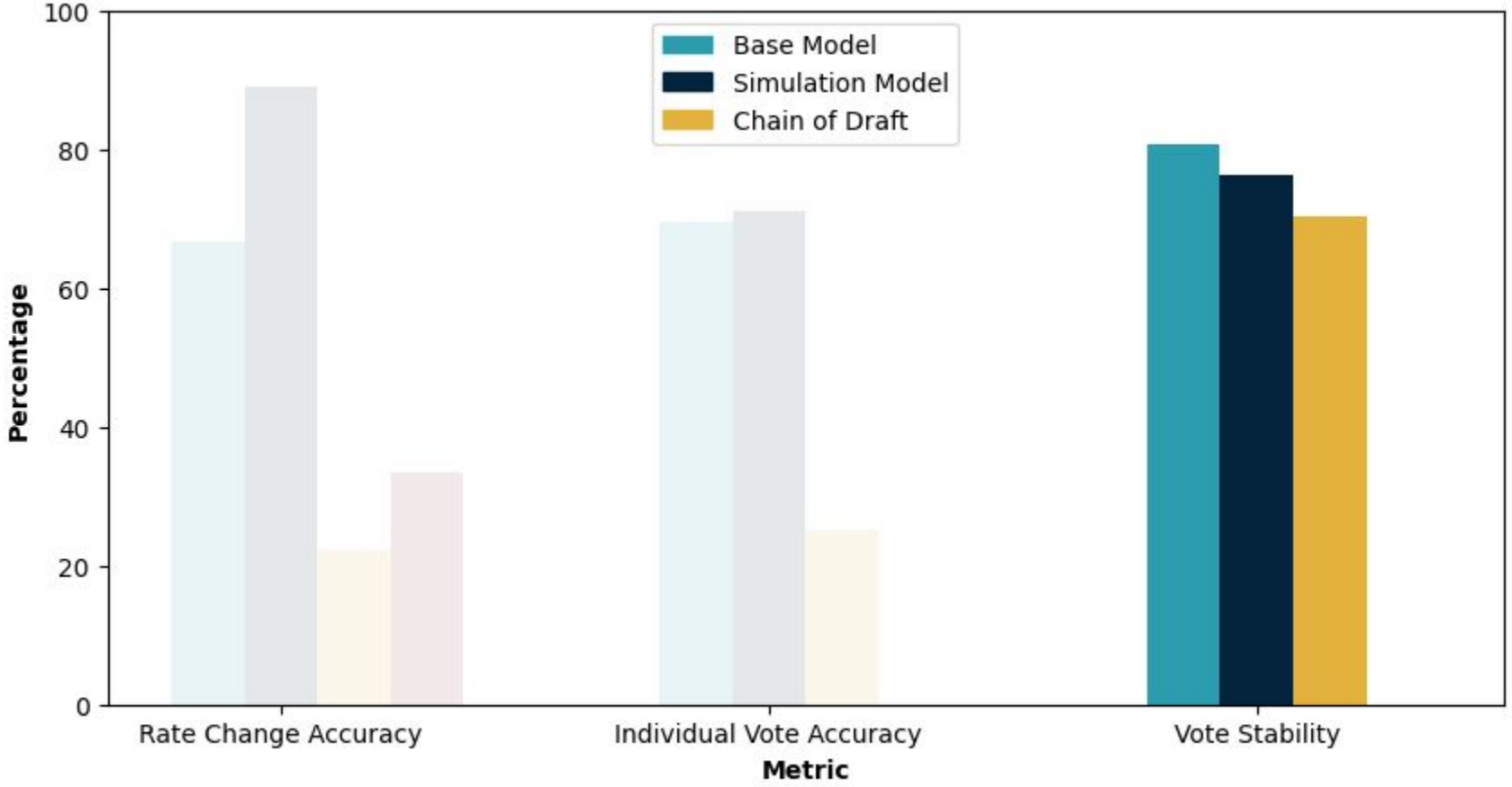
Results

Adding Simulations Improves Voting Accuracy Over Base Model



Results

All Multi-Agent Models Are Between 70 - 80% Stable



Results by Meeting

Model	Feb-23	May-23	Jun-23	Jul-23	Mar-24	Sep-24	Nov-24	Dec-24	Mar-25
Real FOMC Decision	+25	+25	0	+25	0	-50	-25	-25	0
Linear Regression	+50	+25	+25	0	0	0	-25	0	-25
Base Multi-Agent	+50	+25	0	+50	0	-50	-50	-25	0
Base + Simulation	+25	+25	0	+50	0	-50	-25	-25	0
Chain of Draft	0	0	0	0	0	0	0	+25	0

Conclusion

What do these Results Mean ?

- Established a multi-agent framework that **reasons and discusses**
- Implemented **Backtesting** design to gauge reliability of multi-agent framework

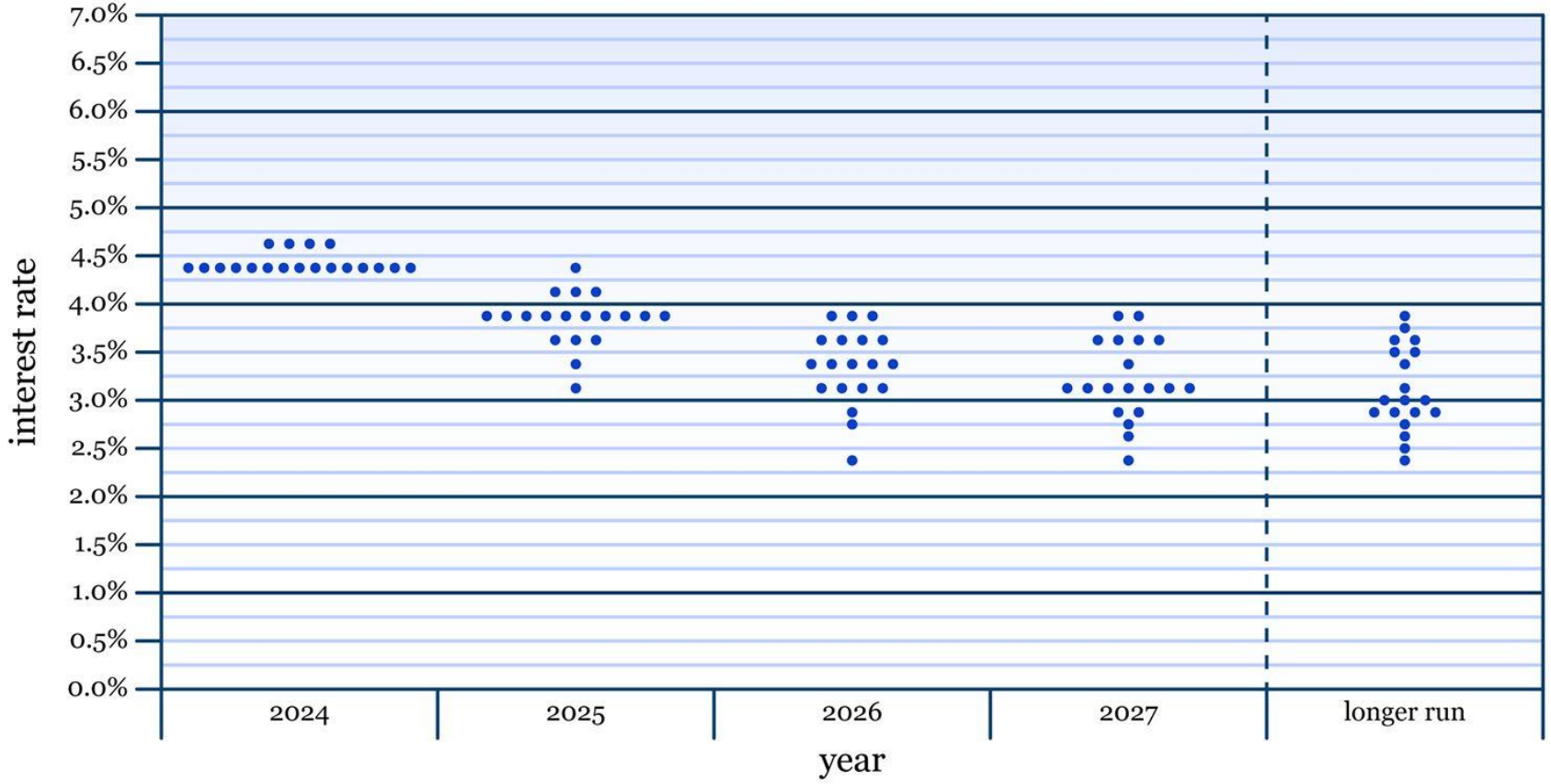
Lays the groundwork for BNY to **integrate** their own model (Eliza), **validate** it using a robust backtesting design, and **incorporate** additional private data

References

1. Li, P., Castelo, N., Katona, Z., & Sarvary, M. (2024). Frontiers: Determining the validity of large language models for automated perceptual analysis. Marketing Science, 43(2), 254–266.
2. Xu, S., Xie, W., Zhao, L., & He, P. (2025). Chain of Draft: Thinking Faster by Writing Less. arXiv preprint arXiv:2502.18600. https://arxiv.org/abs/2502.1860
3. Seok, S., Wen, S., Yang, Q., Feng, J., & Yang, W. (2024). MiniFed: Integrating LLM-based Agentic-Workflow for Simulating FOMC Meeting. arXiv preprint arXiv:2410.18012. https://arxiv.org/abs/2410.18012
4. Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (pp. 1–22).
5. Federal Reserve. (n.d.). FOMC meeting calendars, statements, and minutes. Retrieved February 2, 2025, from https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm
6. CME Group. (n.d.). CME FedWatch Tool. Retrieved February 15, 2025, from https://www.cmegroup.com/markets/interest-rates/cme-fedwatch-tool.html
7. Tadle, R. C. (2022). FOMC minutes sentiments and their impact on financial markets. Journal of Economics and Business, 118, 106021. https://doi.org/10.1016/j.jeconbus.2021.106021
8. Ruman, A. M. (2023). A comparative textual study of FOMC transcripts through inflation peaks. Journal of International Financial Markets, Institutions and Money, 87, 101822. https://doi.org/10.1016/j.intfin.2023.101822


Thank You!
Q&A

Appendix: Fed Dot Plot



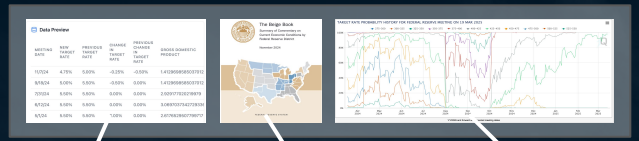
Source: federalreserve.gov

Appendix: Quantitative Predictors

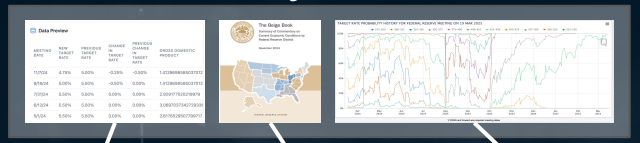
Variable Name	Description
Unemployment Rate	Measures the percentage of the labor force that is jobless and actively seeking employment.
PCE Price Index	Personal Consumption Expenditures index; a key inflation gauge favored by the Fed.
CPI	Consumer Price Index
MICH	University of Michigan's Inflation Expectation; reflects consumer expectations of future inflation.
TB3M	3-Month Treasury Bill yield; short-term government interest rate and a proxy for market expectations.
TB6M	6-Month Treasury Bill yield; provides insight into short-term rate expectations.
M2	Money supply metric that includes cash, checking deposits, and easily convertible near money.
Brave-Butters-Kelley GDP	High-frequency estimate of real GDP growth capturing real-time economic activity.
VIX	Volatility Index; measures market expectations of near-term volatility (often called the "fear index").
Fed Chair	Identifies the current Chair of the Federal Reserve, which can influence policy direction.
Party controls the White House	Indicates which political party holds the presidency, affecting fiscal and economic policy.
Recession Indicator	Signals whether the US is currently in a recession; affects Fed's stance on monetary policy.
Rate Before Meeting	Federal Funds Target Rate in place before the current FOMC meeting.
Value of Change at Prior Meeting	Magnitude of the interest rate adjustments

Model 3: Sequential CoD Architecture

Economist analysis of each source



Member's analysis of each source



Each source is looked at in isolation by both economist and each member before options and discussion

Economist Generates Options

Clustered Member Agents

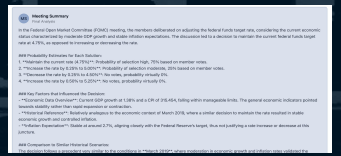
Consolidated Analysis

3-Member discussion

Final Voting Round

Summary Analyst

Results



Model 1: Sequential Architecture “Base”

Input Data

Analyst

Analysis of Implied Probabilities

Economist

Generate 3 Potential Solutions

Clustered Member Agents

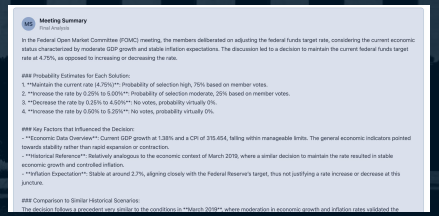
Analysis

Discussion

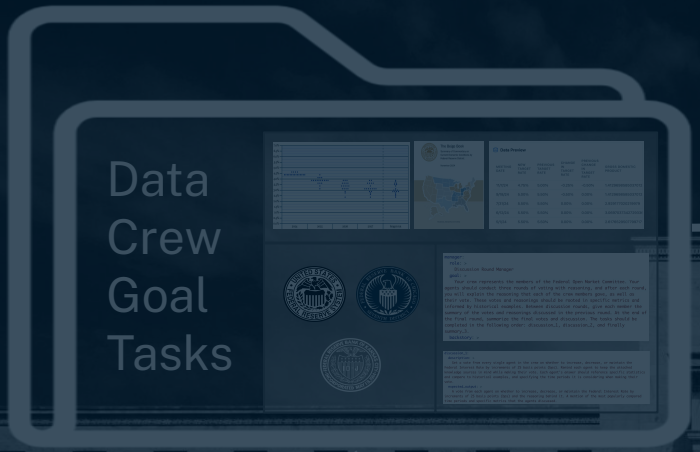
Voting

Summary Analyst

Results



Simulation Process-Hierarchical



Discussion Manager

Crew Discussion

Reveal and Summarize

Crew Reflection