AI-simulated CFO Sentiment

STEVE LECOMPTE | JULY 7, 2026 | POSTED IN: **INVESTING EXPERTISE**, **SENTIMENT INDICATORS**

Can large language models (LLM) simulate market-scale, real-time business sentiment from the Chief Financial Officer (CFO) perspective? In their June 2026 paper entitled "CFOs Meet LLMs", John Graham, Campbell Harvey and Manish Jha use an LLM (GPT-5.4) to simulate CFOs of specific firms. Simulations involve two prompts for each of 6,075 actual responses to the quarterly Duke-Federal Reserve CFO Survey during 2002 through 2025:

1. System prompt – establishes context, persona and analytical instructions, including instructions to collect from the web CFO interviews and analyst reports, price targets and consensus revenue estimates. Instructions include a temporal restriction to ensure that information collected was available before the matched survey response date.

2. User prompt – provides firm profile (industry, revenue, headcount and geographic footprint), respondent history (prior CFO optimism scores when available and the LLM's own prior forecasts) and the key survey question: "Rate your optimism about the overall U.S. economy on a scale from 0–100, with 0 being the least optimistic and 100 being the most optimistic." Again, a temporal restriction imposes an information cutoff date matching the actual survey date.

They run each pair of prompts three times a few seconds apart and average outputs to suppress LLM statistical variation. Finally they match every LLM average output to the response of the actual CFO at the same firm in the same quarter. Using past CFO Survey results and associated firm data spanning 2002 through 2025, *they find that:*

- Overall, the LLM does a good job of mimicking CFOs.

    - Average LLM optimism is 56.6, somewhat lower than the 60.1 for actual CFO survey responses. The two distributions have comparable dispersions across individuals.
    - LLM optimism scores significantly predict actual individual CFO responses.
    - A regression of individual CFO responses versus matched LLM scores generate a coefficient more than six standard errors from zero, surviving firm and year-quarter fixed effects.

- LLM predictive accuracy:
    - Persists after controlling for prior-quarter CFO responses, confirming that the LLM is not simply echoing the last iteration.
    - Increases with both the amount of: (1) firm information provided; and, (2) the individual CFO history available. For the latter, R-squared ranges from near 0.10 for no prior history to nearly 0.50 when more than a handful of prior survey responses are available.
    - Persists after controlling for the Michigan Consumer Sentiment Index and the Survey of Professional Forecasters, indicating that the model captures incremental executive sentiment.

In summary, *evidence suggests that LLMs may offer credible high-frequency estimates of business executive sentiment for financial research*.

Cautions regarding findings include:

- The study requires access to proprietary survey information.
- There may be material ongoing costs of repeated simulations.
- The temporal cutoff restriction imposed in the LLM prompts may not fully protect against look-ahead bias. See "Lookahead Bias in Large Language Model Training Data".
- Business sentiment may not be a useful predictor of asset returns. See "CFO U.S. Economic Sentiment and Stock Market Returns".

See also "Mimicking Economic Expertise with LLMs".

Further Reading