RESEARCH

Cloning CFOs

Digital twins are computer models that mimic the real world. Advances in AI and compu power are making it possible to model individual executives.

MATT ROBINSON
SEP 08, 2026 • PAID

Hey, I'm Matt. I'm a former Bloomberg News reporter, and you're reading AI Street, where report on how Wall Street uses AI. On Tuesdays, I highlight novel research, emerging use ca and expert interviews.

A big challenge in measuring how well AI predicts outcomes is look-ahead bias: quizzing a model on a test it's already taken.

AI providers generally don't disclose their training data, so researchers are in the d on whether a model has already seen the outcome they are asking it to predict.

But if the data has *never* been made public, you can be more confident that the mod is actually making a prediction.

Duke has one such dataset. Its Fuqua School of Business has surveyed CFOs quarte since 1996, with the Atlanta and Richmond Feds joining the project in 2020. The survey's CFO Optimism Index has historically been a leading indicator of hiring an GDP growth, while research using the responses has also found that CFO expectati predict corporate investment.

The aggregate results are published. The individual executives' answers remain confidential.

In a new paper, CFOs Meet LLMs, Duke finance professors John Graham and Campbell Harvey teamed up with Georgia State University professor Manish Jha to test whether an LLM could act as a digital twin for an individual executive.

The test: Could an LLM predict how a particular CFO answered in a particular quarter?

The researchers analyzed 6,075 survey responses from public-company CFOs betwe 2002 and 2025. In each response, the executive rated their optimism about the overa US economy on a scale from zero to 100.

GPT-5.4 was given the executive's real name, company and ticker, along with detail such as revenue, employment, foreign exposure and credit rating. It could search fo company information, analyst reports and economic data available at the time, but told to ignore anything published after the survey date. If a CFO had taken the surv before, the model also got up to 12 of their earlier confidential optimism scores.

What is a digital twin?

A digital twin is a computer model of something real, loaded with data so it can mimic how the original behaves. A digital twin of a jet engine might use temperature and vibration readings to predict when a component will fail. Bologna is building a twin of the entire city to test how changes in transportation, energy use and flood protection might play out.

The idea dates to NASA's Apollo-era simulators. More data and computing power now make it possible to build twins of systems as complicated as cities, supply chains and, in this paper, a CFO.

Insurance companies are beginning to explore digital twins, according to Samarth Swarup, a University of Virginia researcher who builds digital twins

of societies.

"You want to know: How many claims can we expect? Which households are more at risk depending upon how evacuation happens? And what can we do proactively to reduce that?" Swarup told me.

Results

The researchers ran each prompt three times and averaged the results. The model g very similar scores each time, with correlations of about 0.96 across the runs.

Compared with the real CFOs, the model was slightly more pessimistic and made t executives look more alike than they really were, producing fewer very high or very scores.

Across all responses, the synthetic scores explained 27% of the differences in CFO optimism. The relationship held after the researchers accounted for each executive' usual level of optimism, conditions during the quarter and the CFO's most recent answer.

"It is not asking whether the model can separate persistent optimists from persiste pessimists," Jha said in an email. "It is asking whether it can tell when a particular CFO turns more or less optimistic than their own usual level."

Related Research

Can AI Simulate How Investors Trade News?

MATT ROBINSON • APR 21

Read full story →

AI Replicates Human Investor Biases

MATT ROBINSON • MAR 11

Read full story →

The researchers accounted for persistent differences among firms, each CFO's mos recent answer and whatever was moving everyone's mood that quarter. The model's score remained statistically significant.

But it worked much better when more history was available. Among responses with previous answers, the model explained 10% of the variation. That rose to 33% when one to three previous answers were available and 49% when more than three were available. The model was much better at extending a known executive's past views than building one from public information alone.

"The surprising part for me was how much the conditioning did, more than the raw capability of the model," Jha said in an email.

The model isn't a stand-in for a human CFO. It does not have access to the private information that makes an executive's view valuable: current orders, internal costs, financing discussions and changes in corporate strategy. The authors described the synthetic scores as a possible supplement to human surveys, not a replacement.

The study also reconstructed past responses from 2002 through 2025 rather than making new predictions in real time. The researchers plan to test the system agains future survey rounds.

For now, the system looks more useful for extending an existing panel of known respondents than for creating thousands of synthetic CFOs from scratch.

TL;DR

GPT-5.4 could predict how individual CFOs rated the economy, but it worked far better when given their previous survey scores.

Behind the paywall

Further commentary from Manish Jha on where digital twins are useful in financ research and policy — and why he wouldn't trade on one yet.

← Previous