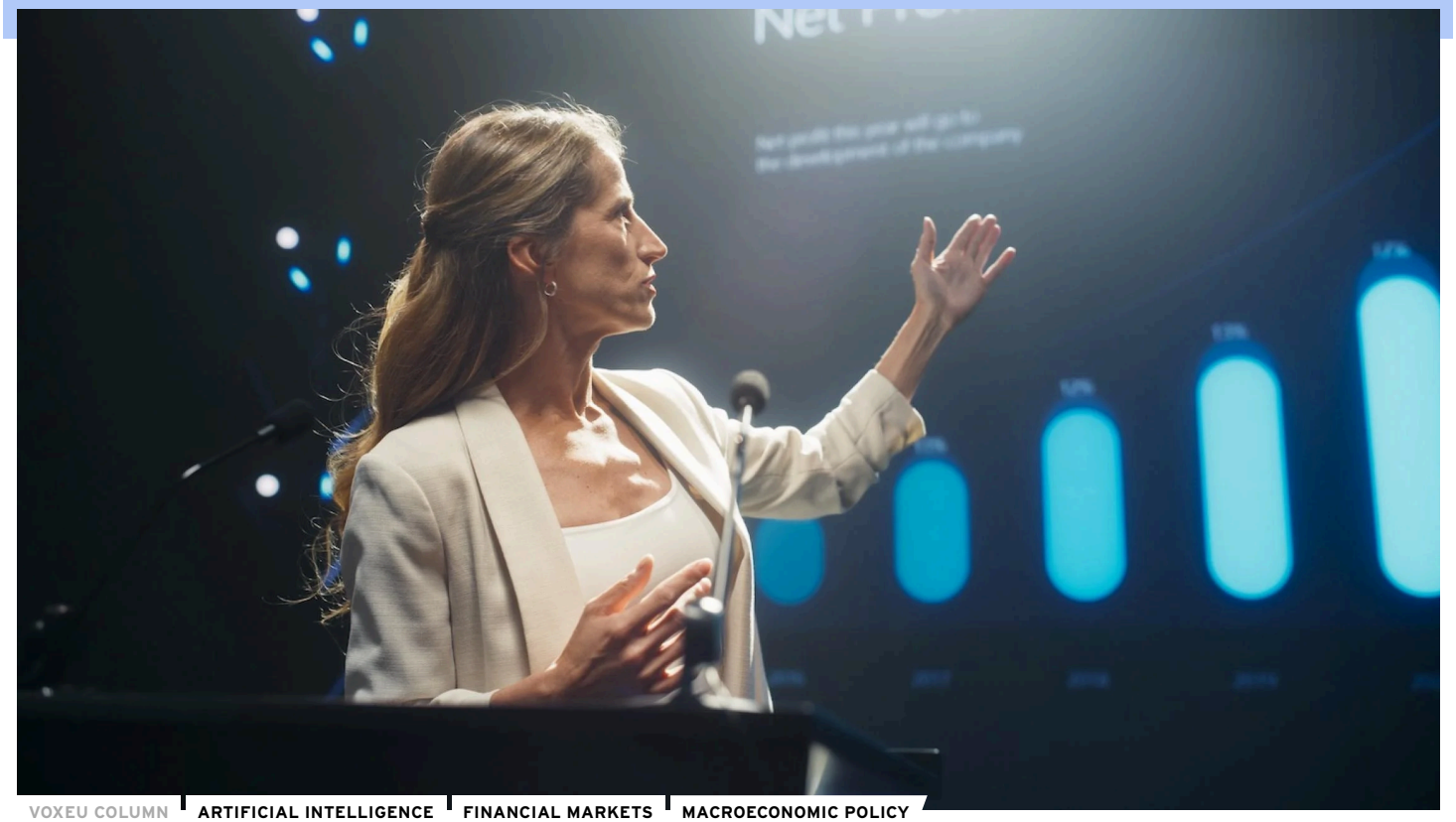


VOXEU COLUMN | ARTIFICIAL INTELLIGENCE | FINANCIAL MARKETS | MACROECONOMIC POLICY

Reading the economy from earnings calls with generative AI

Manish Jha, Jialin Qian, Michael Weber, Baozhong Yang / 12 Nov 2026

Official statistics are getting harder to collect and can arrive late. This column uses generative AI to read managers' economic outlooks from more than 120,000 corporate earnings calls in the US. The resulting score predicts GDP, employment, and wages beyond the next quarter and carries information that professional forecasters do not.

AUTHORS



Manish Jha



Jialin Qian



Michael Weber



Baozhong Yang

In October 2025, a 43-day US government shutdown delayed the September jobs report, and the Federal Reserve set policy with much of its usual data missing. Analysts turned to private indicators such as purchasing managers' surveys and job-postings trackers (J.P. Morgan 2025). The concern outlasted the shutdown. A June 2026 review by the US Government Accountability Office warns that falling survey response rates are a risk to the quality of the jobs report (GAO 2026). The problem is not only American: in the UK, falling response rates led the Office for National Statistics to suspend its headline Labour Force Survey estimates in late 2023. This raises an old question: what timely, forward-looking information about the economy exists outside the official statistics?

Firms are an obvious place to look. Listed US firms employ close to 30% of non-farm workers and account for more than a quarter of GDP (Schlingemann and Stulz 2022). Firms' expectations shape their decisions (Coibion et al. 2018, Candia et al. 2023), and columns on VoxEU have studied how firms react to news (Born et al. 2023) and how they formed expectations during the recent inflation surge (Gautier et al. 2025). Most of this evidence comes from surveys, which are informative but costly to run and hard to scale.

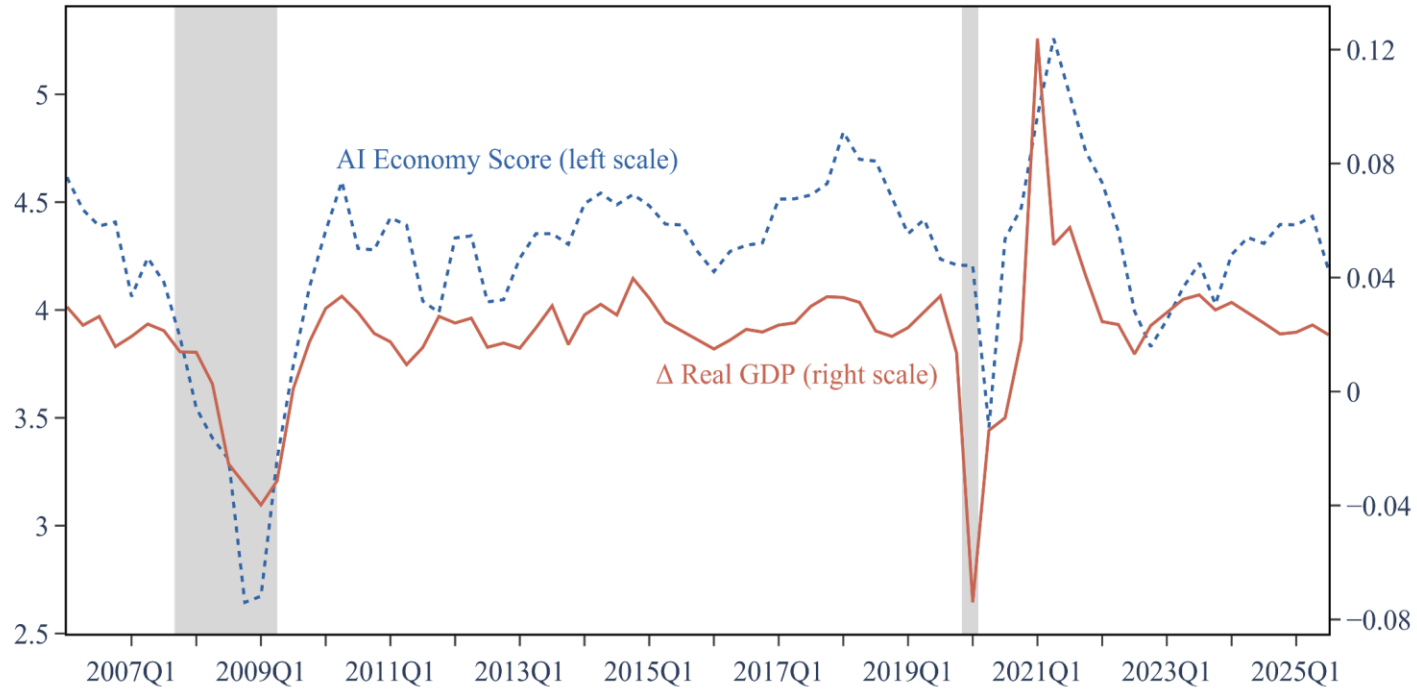
Reading managers' outlook from what they already say

In a recent paper (Jha et al. 2026), we take a different route. In quarterly earnings calls, managers of US public firms talk about their own results, but also about their customers, input costs, hiring, tariffs, and interest rates. We use a generative AI model (OpenAI's GPT-5 mini) to read 125,830 transcripts from 5,513 firms between 2006 and 2025. Our approach adds to a growing literature that measures beliefs from text with large language models (Bybee 2025; see Ash and Hansen 2023 for an overview on VoxEU).

The model first keeps only the sentences in which management speaks about the US economy. It then rates how management expects US conditions to change in the next quarter, from 1 (severe deterioration) to 7 (strong expansion). Averaging across firms gives a national measure, which we call the AI Economy Score. Averaging within a sector gives an industry score.

The model's reasoning lines up with economic intuition. Low scores are explained by weak consumer demand, supply chain disruption, and tariff-driven uncertainty. High scores are explained by strong demand and tight labour markets. Figure 1 shows that the national score tracks realised GDP growth closely and turned down ahead of the 2008 financial crisis. It did not anticipate the collapse in early 2020, which is unsurprising given how little warning the pandemic gave.

Figure 1 AI Economy Score and realised real GDP growth, 2006–2025



***Notes*: The AI Economy Score (left scale) is the quarterly average across firms of the 1–7 score of managers' expectations for the US economy in the next quarter. Realised GDP growth (right scale; 0.04 = 4%) is the percent change in real GDP in quarter t+1 relative to quarter t−3. Shaded areas are NBER recessions. Source: Jha et al. (2026).**

What the score predicts

The AI Economy Score predicts next-quarter real GDP growth after controlling for the term spread, the real federal funds rate, the credit spread of Gilchrist and Zakrajšek (2012), and past GDP growth. A one standard deviation increase in the score is associated with 1.3 percentage points higher real GDP growth. The score predicts GDP and employment growth up to five quarters ahead and wage growth for about ten quarters. For employment and wages, it outperforms the credit spread, which is plausible since managers speak directly about hiring and pay.

The score also holds up against professional forecasters. In a horse race with the median GDP forecast from the Survey of Professional Forecasters, the two have similar explanatory power, and each remains significant when both are included. Managers and professional forecasters therefore appear to hold partly different information.

Beyond the national average

Because the measure starts at the firm level, it can be built for any sector or group of firms. The industry score predicts sectoral real GDP growth up to four years ahead. A manager's view of the economy relative to sector peers predicts the firm's own sales, employment, and wages for several years. In other words, the cross-section of managerial views contains information that the national average hides.

We also ask the model about 13 other topics, including demand, output prices, input costs, hiring, investment, and the cost of capital. Combining the 14 scores, with weights estimated only from past data, extends the horizon over which the national measure predicts GDP growth from five to eight quarters.

Information or sentiment?

Why do managers' expectations predict the economy? One possibility is that managers see orders, prices, and hiring conditions before they show up in official data. Alternatively, their optimism may be sentiment that moves spending regardless of fundamentals, the noise channel discussed by Barsky and Sims (2012), Blanchard et al. (2013), and Chahrour and Jurado (2018). We find that the predictive power comes from managers whose past views of the economy were accurate or unbiased, while scores from historically inaccurate managers predict little. Managers of more focused firms are also more informative than those of diversified firms, consistent with deep knowledge of one business giving a clearer signal. This pattern is more consistent with information than with noise.

Can we trust a language model with history?

A natural concern is look-ahead bias, since the model was trained on text written after most of the calls in our sample (Glasserman and Lin 2023). We address this in three ways. First, following Engelberg et al. (2025), we strip transcripts of firm names, dates, and other identifying details, and the score still predicts GDP growth. Second, firm- and industry-level scores built only from calls after the model's training cutoff still predict sales and sectoral output. Third, Google's open-source Gemma model gives similar results. The score also remains significant after controlling for dictionary-based sentiment in the calls, so it is not simply picking up upbeat language.

Implications for policy

We do not see this measure as a substitute for official statistics or surveys. Earnings calls cover only listed firms, and surveys can ask precise questions. But the measure is built from public text available within weeks of each quarter's end, so it does not depend on the functioning of a statistical agency or on survey response rates. It is cheap to produce, including with open-source models that a central bank can run in-house. And it can be cut by sector, which gives an early read on where a slowdown or a pickup is concentrated.

Central banks already collect anecdotal reports from businesses, for example in the Federal Reserve's Beige Book, the Bank of England's Agents' summaries, and the ECB's regular contacts with non-financial companies. Our results suggest that a systematic version of this exercise, built from what managers say in public, carries information about the economy well beyond the next quarter.

References

Ash, E and S Hansen (2023), "Large language models for economic research: Four key questions", VoxEU.org, 27 July.

Barsky, R B and E R Sims (2012), "Information, animal spirits, and the meaning of innovations in consumer confidence", *American Economic Review* 102(4): 1343–77.

Blanchard, O J, J-P L'Huillier and G Lorenzoni (2013), "News, noise, and fluctuations: An empirical exploration", *American Economic Review* 103(7): 3220–57.

Born, B, Z Enders, M Menkhoff, G J Müller and K Niemann (2023), "Island illusion: How firms react to news", VoxEU.org, 11 January.

Bybee, J L (2025), "The ghost in the machine: Generating beliefs with large language models", arXiv:2305.02823.

Candia, B, O Coibion and Y Gorodnichenko (2023), "The macroeconomic expectations of firms", in R Bachmann, G Topa and W van der Klaauw (eds), *Handbook of Economic Expectations,* Elsevier: 321–53.

Chahrour, R and K Jurado (2018), "News or noise? The missing link", *American Economic Review* 108(7): 1702–36.

Coibion, O, Y Gorodnichenko and S Kumar (2018), "How do firms form their expectations? New survey evidence", *American Economic Review* 108(9): 2671–713.

Engelberg, J, A Manela, W Mullins and L Vulicevic (2025), "Entity neutering", SSRN working paper.

GAO – US Government Accountability Office (2026), "Federal statistics: Stakeholders said jobs report generally meets their needs, but opportunities exist to improve data quality", GAO-26-107538, 2 June.

Gautier, E, F Savignac and O Coibion (2025), "How French firms navigated the inflation surge: Lessons for expectations and decision-making", VoxEU.org, 3 June.

Gilchrist, S and E Zakrajšek (2012), "Credit spreads and business cycle fluctuations", *American Economic Review* 102(4): 1692–720.

Glasserman, P and C Lin (2023), "Assessing look-ahead bias in stock return predictions generated by GPT sentiment analysis", *Journal of Financial Data Science.*

J.P. Morgan (2025), "The September jobs report is delayed by the government shutdown", J.P. Morgan Insights, 3 October.

Jha, M, J Qian, M Weber and B Yang (2026), "Generative AI, managerial expectations, and economic activity", NBER Working Paper 35949.

Schlingemann, F P and R M Stulz (2022), "Have exchange-listed firms become less important for the economy?", *Journal of Financial Economics 143*(2): 927–58.

AUTHORS

Manish Jha
Assistant Professor of Finance, Georgia State University

Jialin Qian
Assistant Professor of Finance, Rochester Institute Of Technology

Michael Weber
Professor of Finance, Alexander von Humboldt Professor, and head of the European Expectations Center, ESMT Berlin

Baozhong Yang
H. Talmage Dobbs Jr Chair in Finance and Associate Professor of Finance, Georgia State University

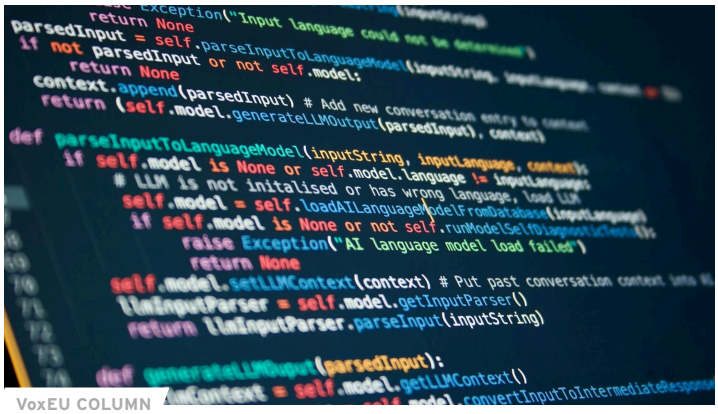
THEMES

ARTIFICIAL INTELLIGENCE FINANCIAL MARKETS MACROECONOMIC POLICY

KEYWORDS

EARNINGS CALLS AI ECONOMIC STATISTICS

SHARE



VoxEU COLUMN

Large language models for economic research: Four key questions

Elliott Ash, Stephen Hansen

27 JUL 2023 FRONTIERS OF ECONOMIC RESEARCH



VoxEU COLUMN

What corporate earnings calls reveal about the AI stock rally

Michele Ca' Zorzi, Gianluigi Lopardo, Ana-Simona Manu

28 SEP 2025 FINANCIAL MARKETS



VoxEU COLUMN

Leveraging large language models for large-scale information retrieval in economics

Prashant Garg, Thiemo Fetzer

19 FEB 2025 ARTIFICIAL INTELLIGENCE / FRONTIERS OF ECONOMIC RESEARCH

THEMES & CURRENT ISSUES

Artificial Intelligence
Business Cycles
Central Banking
Climate Change
Competition Policy
COVID-19
Development & Growth
Economic History
Energy
EU Economic Architecture
Exchange Rates
Finance & Fintech

Financial Markets
Financial Regulation and Banking
Health Economics
Inflation
International Finance
International Trade
Labour Markets
Migration
Monetary Policy
Political Economy
Poverty and Income Inequality
Productivity and Innovation
Taxation
Ukraine Initiative
Women in Economics

SIGN UP TO OUR NEWSLETTER

Your email address

SIGN UP

FOLLOW US

Contact Us Cookies Privacy Policy RSS Feeds

Copyright 2026 CEPR /designbysoapbox.com