# Harnessing Generative AI for Economic Insights

Manish Jha[1] Jialin Qian[1]
Michael Weber[2] Baozhong Yang[1]

[1]Georgia State

[2]Purdue University, NBER

SI 2025 Forecasting & Empirical Methods
AI and ML Methods in Economics and Economic Forecasting

# Motivation

**Managers of public corporations have a disproportionate influence on the economy.**

- Opinions are often heeded by the public and the government
- Featured prominently in the media.

# Surveys are Costly

Examples of manager surveys:

- CFO survey, conducted by Duke University (200-400 firms)
- Survey of Firms Inflation Expectations (SoFIE), conducted by Federal Researve Bank of Cleveland (200-400 firms)

**Managerial surveys are costly to conduct.**

- These surveys are at a smaller scale.
- The underlying data are private.

# What do we do

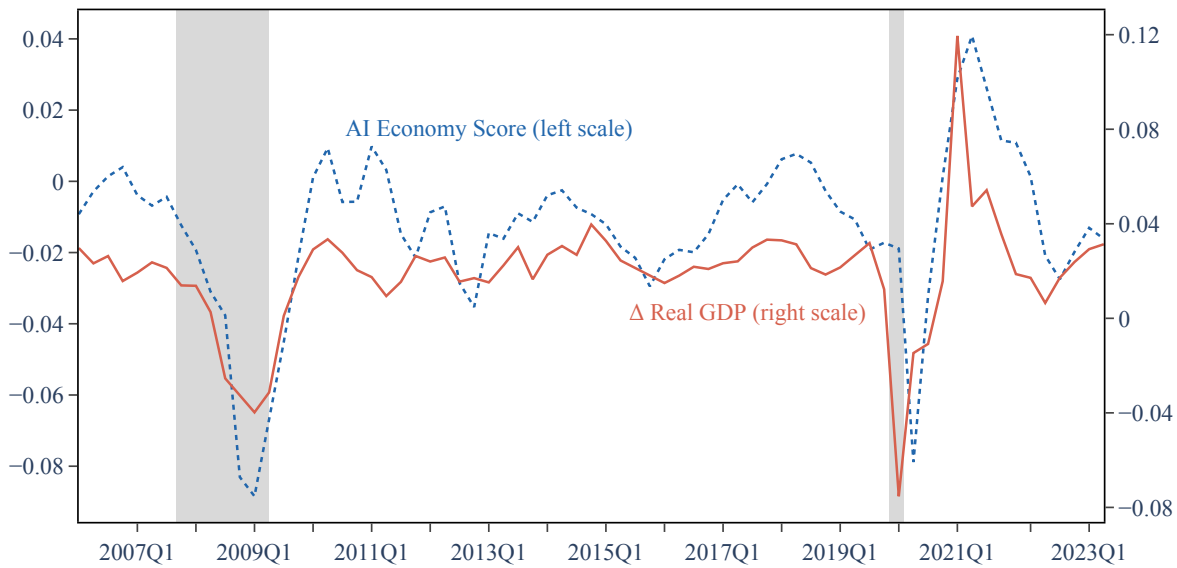
1. Apply Generative AI to conference call transcripts
2. Aggregate quarterly at the national level
3. Obtain managerial expectations about economy - **AI Economy Score**

Our method:

- Covers 5,513 firms
- Sourced from publically available dataset, accessible
- Applicable to other dimensions of interests
- Available at micro-level

# AI Economy Score

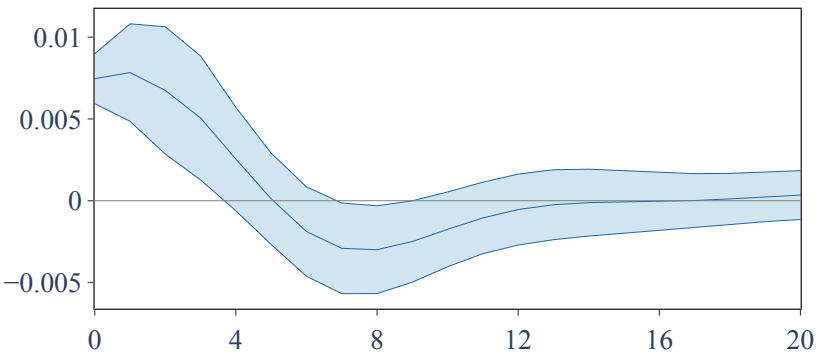
- **Captures managerial expectation** for the US economy in the next quarter.
- Predictive of the next quarter's real GDP
- Predictive power persists in the longer term and for other economic activities - industrial production, employment, and wages



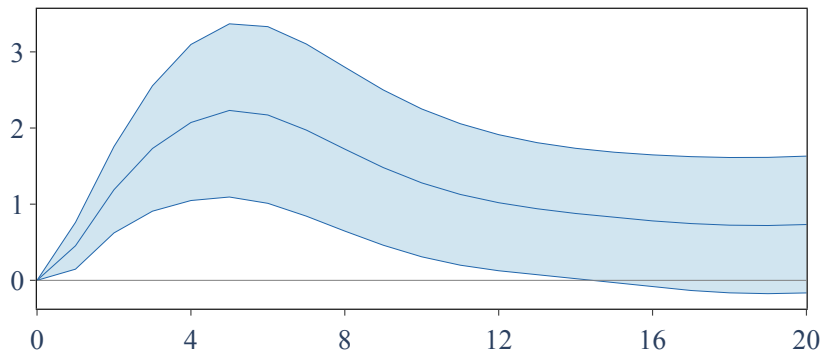
# But these variables affect each other

We employ a vector autoregression framework (VAR)

- Shocks to the AI Economy Score orthogonal to current economic activities influence future activities significantly.
- Higher consumption, investment, and output growth




(e) AI Economy Score




(f) Output growth (cumulative)

# Using More Dimensions

We aso extract 14 different questions on managers' expectations, covering firm financial prospects and macroeconomic trends.

1. **AI Weighted Score** constructed using firm-quarter regression model to predict future firm sales.
2. This score significantly improves GDP growth predictions at national and industry levels.
3. Outperforms simpler AI-based measures, surpassing the original AI Economy Score.
4. Most influential factors: revenue, production, wages, employment, industry prospects, and capital expenditures.

# Contributions 1/2

Economic expectations and real outcomes:

- News about future economic conditions impact business cycles and stability (Barsky and Sims, 2011; Chahrour and Jurado, 2018; Schmitt-Grohé and Uribe, 2012; Blanchard et. al, 2013)
- Changes in consumer confidence and firm expectations influence economic activities (Barsky and Sims, 2012; Coibion, Gorodnichenko, and Kumar, 2018; Coibion et. al, 2022)
- Survey data on subjective beliefs and economic activities (Bhandari, Borovicka, and Ho, 2022; Coibion, Gorodnichenko, and Ropele, 2020; Weber, D'Acunto, Gorodnichenko, and Coibion, 2022)

Our method complements standard surveys. Managerial expectations are strongly predictive of future economic activities.

# Contributions 2/2

- Variables indicative of economic activities
  - Our approach generates variables with long-run predictability: aggregate, industry-level, state-level, and firm-level expectations – helpful at both the macro and the micro-levels.
  - Corporate bond premia (Gilchrist and Zakrajšek, 2012)
  - Expectations from WSJ news articles (Bybee, 2023)
  - National sentiments based on book corpora (Jha, Liu, and Manela, 2022) or local newspapers (van Binsbergen et al., 2024).
- Use of generative AI in finance and economics
  - Fast-growing literature, e.g. Hansen and Kazinnik (2023); Jha, Qian, Weber, and Yang (2023); Kim, Muhn, and Nikolaev (2023, 2024); Li, Mai, Shen, Yang, and Zhang (2023); Yang (2023); Ouyang, Yun, and Zheng (2024).
  - Sheng, Sun, Yang, and Zhang (2024) build on our approach to investigate the reliance on GenAI by investment companies.

# Data and Methodology

# Data Sources

- Conference Call Transcripts
  - **SeekingAlpha** for the period 2006 to 2020.
  - **FinancialModelingPrep** for the period 2021 to 2023.
- Real Outcomes
  - **Federal Reserve Economic Data**: GDP, Industrial Production, Employment, Wages, Federal Funds Rates and Treasury Yields
  - **U.S. Bureau of Economic Analysis**: Industry-level economic indicators.
  - **Compustat** and **CRSP**: Firm-level variables.
- Survey Forecast Data
  - **Survey of Professional Forecasters website**: GDP Forecast, Industrial Production Forecast, Employment Forecast

# Economic Target Variables

- Macro-level and Industry-Level indicators
  - Real GDP; Industrial Production; Employment; Wages
- Firm-level indicators
  - Value-Added; Sales; Employment; Wages
- Macro Control Variables
  - Term Spread: 10Y - 3M yield
  - Real FFR: real federal funds rates
  - GZ Spread: an indicator of overall risk sentiment in the credit markets (Gilchris and Zakrajsek, 2012)

# Constructing AI Economy Score

We provide the following prompt:

*"The following text is an excerpt from a company's earnings call transcripts. You are a finance expert. Based on this text only, please answer the following question. Over the next quarter, how does the firm anticipate a change in optimism about the US economy? There are five choices: Increase substantially, increase, no change, decrease, and decrease substantially. Please select one of the above five choices for each question and provide a one-sentence explanation of your choice for each question. The format for the answer to each question should be "choice - explanation." If no relevant information is provided related to the question, answer "no information is provided."*
*[Part of an earnings call transcript.]*

# Aggregating at the Conference Call Level

1. We split conference call texts into 2,500-word chunks.
2. We assign a score of -1, -0.5, 0, 0.5, and 1 for each of the given choices (Decrease substantially; Decrease; No change; Increase; Increase substantially), respectively.
3. We take the average of the scores across multiple chunks of one earnings call to obtain a firm-quarter-level measure, AI Firm Score.
4. We take the average of AI Firm Score across 18 industries to obtain a industry-quarter-level measure, AI Ind. Score.
5. We take the average of AI Firm Score for all firms to obtain a macro-quarter-level measure, **AI Economy Score**.

# Summary Statistics

| | N | Mean | SD | p25 | Median | p75 |
|---|---|---|---|---|---|---|
| *AI Economy Score* | 72 | -0.013 | 0.022 | -0.020 | -0.010 | -0.002 |
| *Real GDP* | 72 | 0.009 | 0.019 | 0.006 | 0.012 | 0.015 |
| *Industrial Production* | 72 | 0.001 | 0.034 | -0.006 | 0.010 | 0.018 |
| *Payroll Employment* | 72 | 0.004 | 0.022 | 0.005 | 0.008 | 0.010 |
| *Wage* | 72 | 0.014 | 0.005 | 0.009 | 0.013 | 0.016 |
| *Term Spread* | 72 | 0.014 | 0.012 | 0.005 | 0.016 | 0.023 |
| *Real FFR* | 72 | 0.014 | 0.018 | 0.001 | 0.003 | 0.022 |
| *GZ Spread* | 72 | 0.022 | 0.011 | 0.016 | 0.020 | 0.025 |

- The dependent variables - logarithm of the ratio of the economic indicator in quarter $t + n$ to quarter $t - 1$.
- Median and Mean AI Economy Score is slightly negative

# Validation

# Validation Tests

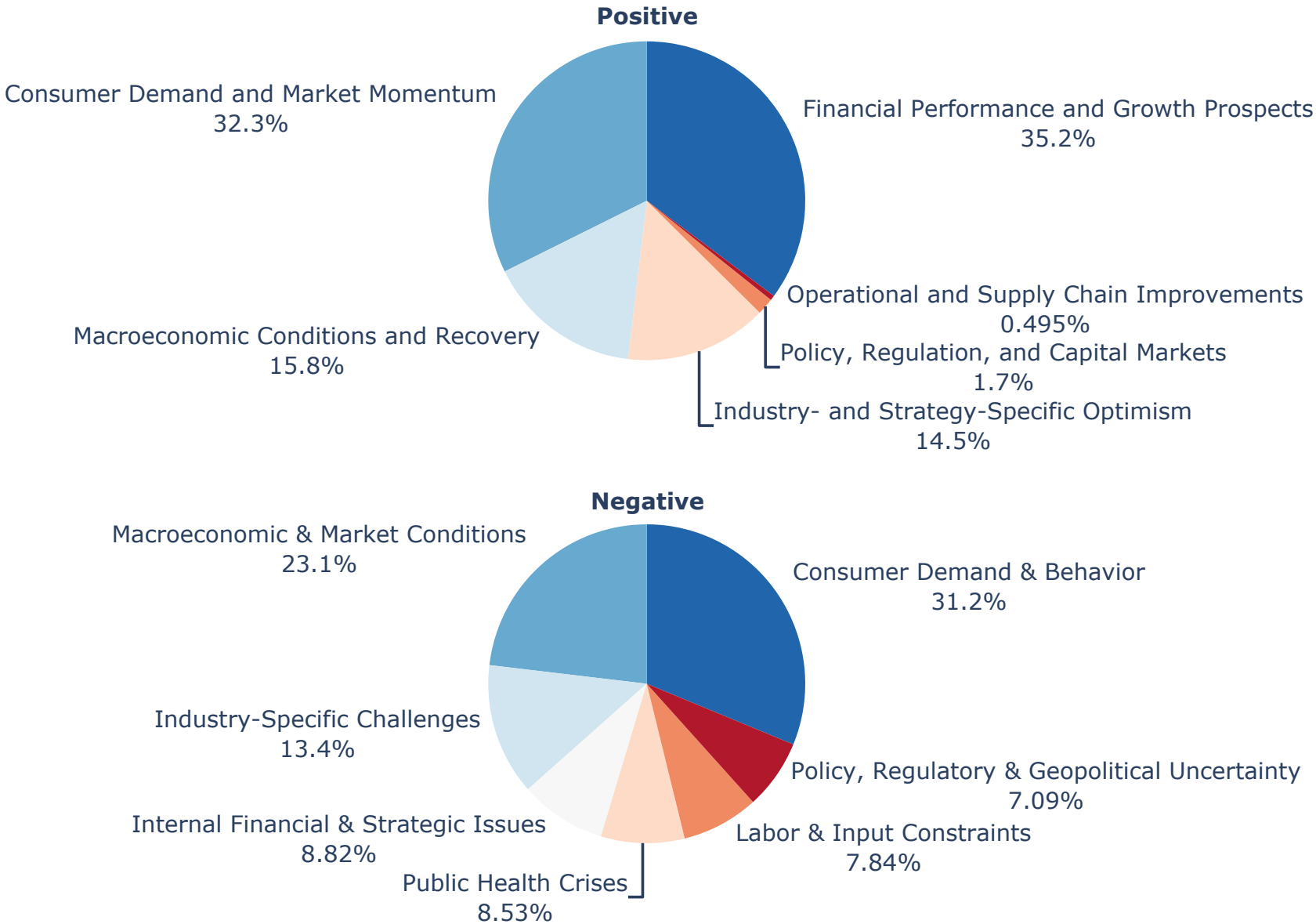
Manual Review and Frequency Analysis

- Go over a sample of the conference calls and responses
- Analyze "explanations"
  - Distinct topics
  - Frequently occurring n-grams in positive/negative scores

Trend Analysis

- Comparison with **actual GDP growth**
- Comparison with SPF forecast (Figure 2 in paper)
- Heterogeneity across industries.

# Explanations for Positive/Negative Scores

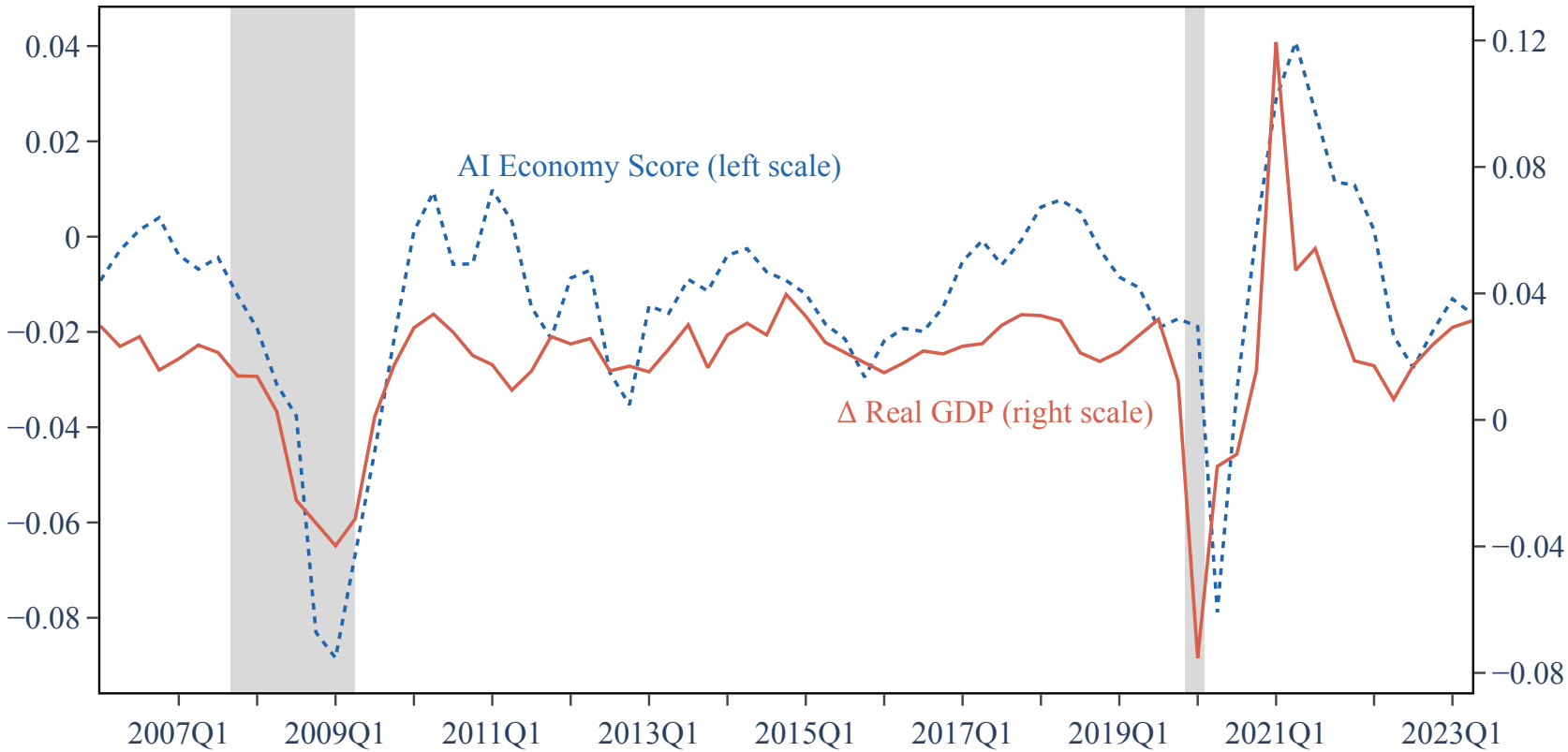


# Phrases Associated with Low and High Scores

Calls with low AI Economy scores often mention global uncertainty, cautious outlooks, and currency risks. In contrast, high scores reflect optimism, strong sales, and positive financial performance.

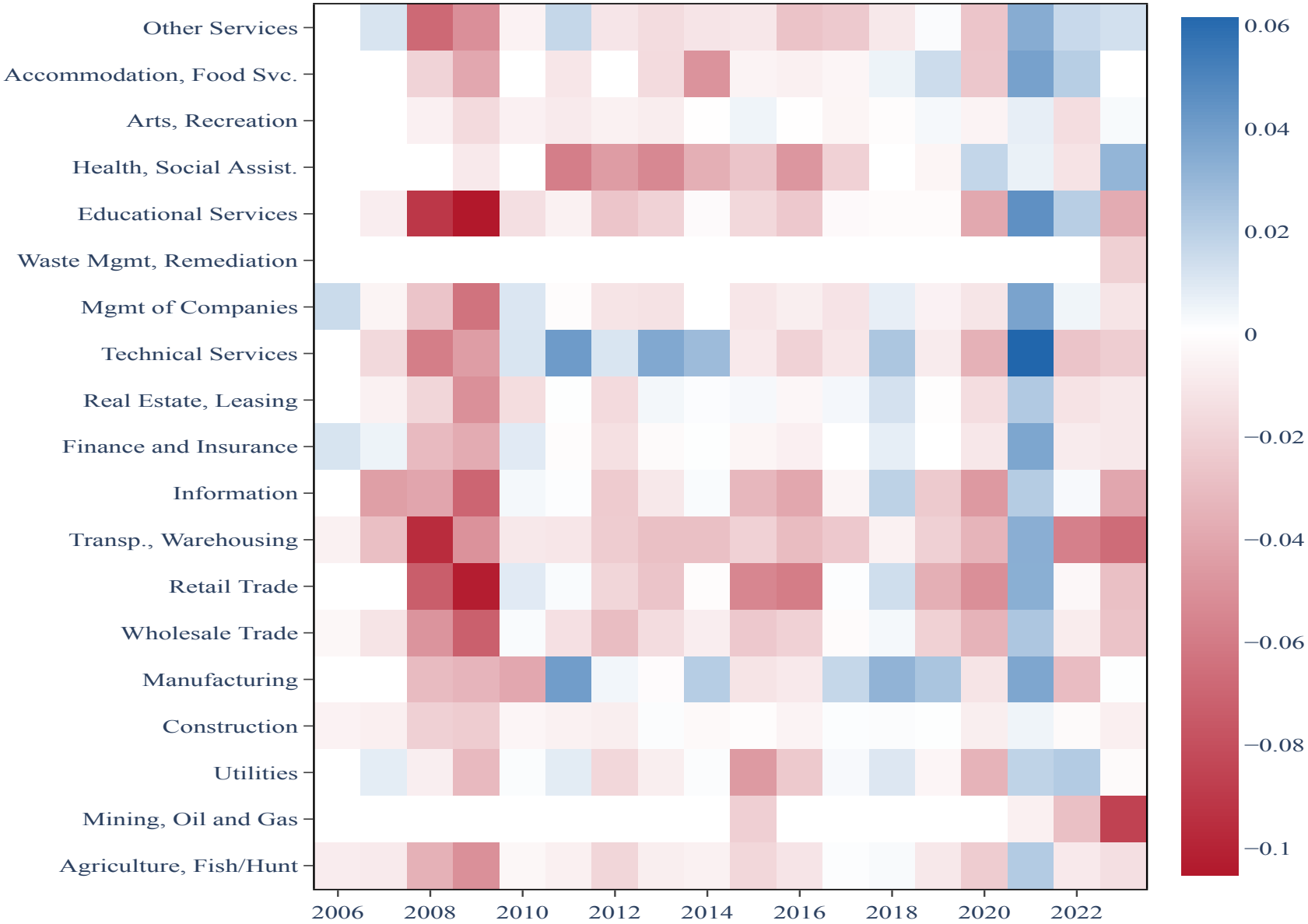
| Low AI Economy Score | High AI Economy Score |
|---|---|
| lower natural gas price | strong first quarter result |
| foreign currency exchange rate | indicate positive economic condition |
| low natural gas price | strong organic revenue growth |
| market condition remain challenge | strong second quarter result |
| would likely decrease optimism | experience strong revenue growth |
| challenge global economic condition | experience strong sale growth |
| global light vehicle production | firm anticipate continue growth |
| retail environment remain challenge | indicate positive market condition |
| business environment remain challenge | strong first quarter performance |
| challenge global business environment | indicate positive business performance |

# AI Economy Score vs. Realized GDP Growth




The AI Economy Score tracks GDP fluctuations well, especially during the 2008 crisis. Its miss in early 2020 is reasonable given the unexpected shock of COVID-19.

# Industry Level - AI Economy Score



# Results

# Results Overview

1. The ability of *AI Economy Score* in predicting realized GDP in the short/long term.
   - 1.1 National Level
   - 1.2 Industry-Level (Table 6-7 in paper)
   - 1.3 Firm-Level (Table 8 in paper)

2. Predicting other economic indicators in the short/long term: Industrial Production, Employment, Wages (Table 4 in paper)
3. Behavioral Biases: Optimists, Realists, and Pessimists
4. Incorporating multi-dimensional information

# Incremental Predictive Power of Measure

$$ln\frac{Y_{t+1}}{Y_{t-1}} = \alpha + \beta_1 \textbf{AI Economy Score}_t + \beta_2 Term\ Spread_t + \beta_3 Real\ FFR_t + \beta_4 GZ\ Spread_t + \sum_{i=1}^{4} ln\frac{Y_{t-i}}{Y_{t-i-1}} + \varepsilon_t$$

where $Y$ represents economic indicators. Our analysis uses *Real GDP* as the primary indicator, while also incorporating *Industrial Production*, *Employment*, and *Wages* as additional metrics.

# National-Level Results

# Realized GDP

Panel A: Next Quarter

| | (1) | (2) | (3) |
|---|---|---|---|
| | *Real GDP:* Next Quarter | | |
| *Term Spread* | 0.371 | -0.284 | 0.108 |
| | (0.82) | (-0.85) | (0.22) |
| *Real FFR* | 0.0493 | -0.157 | -0.0325 |
| | (0.21) | (-0.82) | (-0.14) |
| *GZ Spread* | -1.320*** | | -0.763* |
| | (-4.48) | | (-1.93) |
| *AI Economy Score* | | **0.625*** ** | **0.340*** ** |
| | | (6.83) | (3.18) |
| R-squared | 0.454 | 0.448 | 0.493 |
| Observations | 72 | 72 | 72 |

Predict realized GDP growth for next quarter. The magnitude of the predicting power is comparable to that of *GZ Spread*.
1SD increase = +1.3% GDP

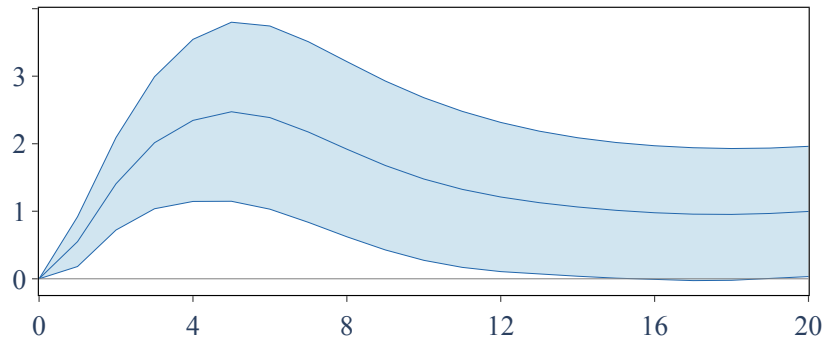
# Realized GDP - longer term

Panel B: Long horizons

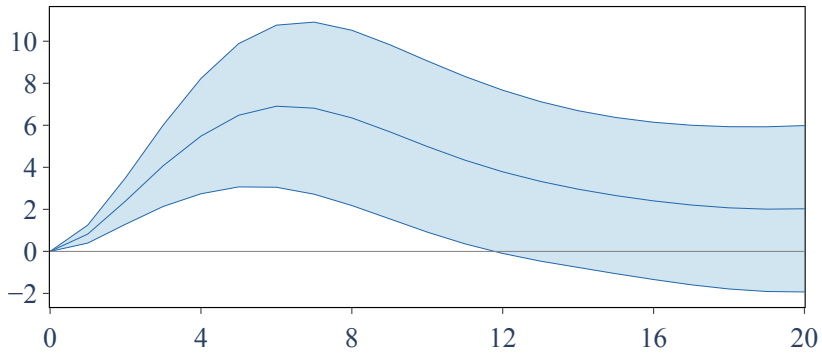
| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | | | *Real GDP* | | |
| | 2 quarters | 3 quarters | 4 quarters | 5 quarters | 6 quarters |
| *Term Spread* | -0.158 | -0.107 | -0.124 | -0.275 | -0.482* |
| | (-0.53) | (-0.29) | (-0.32) | (-0.83) | (-1.86) |
| *Real FFR* | -0.229 | -0.356** | -0.557*** | -0.821*** | -1.103*** |
| | (-1.44) | (-2.03) | (-2.94) | (-3.93) | (-4.82) |
| *GZ Spread* | -0.591** | -0.911*** | -1.206*** | -1.212*** | -1.177*** |
| | (-2.22) | (-3.17) | (-3.31) | (-3.57) | (-3.53) |
| *AI Economy Score* | **0.504*** ** | **0.414*** ** | **0.219*** | 0.178 | 0.0698 |
| | (3.03) | (2.96) | (1.76) | (1.58) | (0.58) |
| R-squared | 0.522 | 0.554 | 0.541 | 0.575 | 0.601 |
| Observations | 71 | 70 | 69 | 68 | 67 |

... persist for up to four quarters.

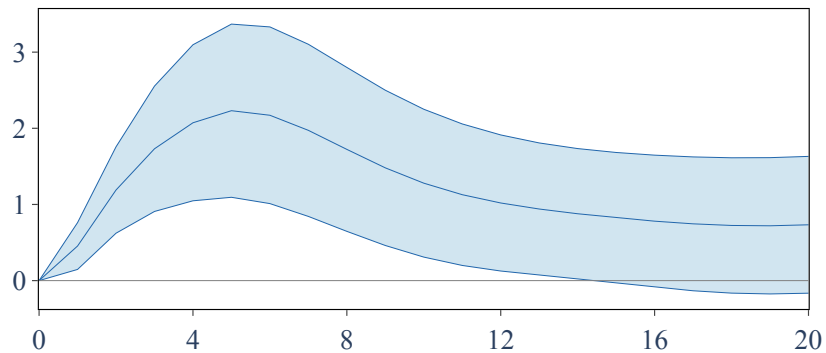
# VAR Analysis of Macroeconomic Implications of Changes in AI Economy Score 1/2



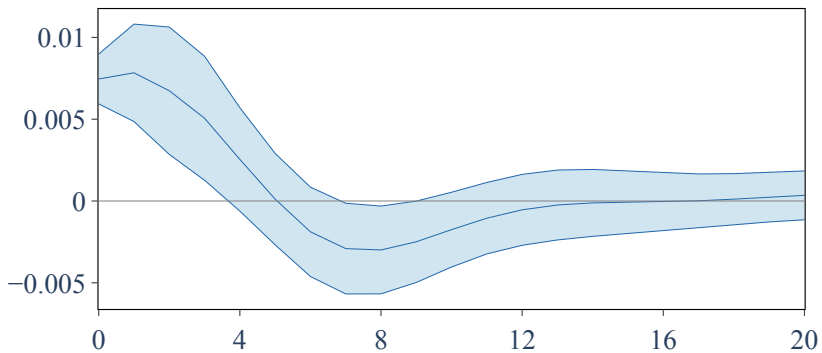

(a) Consumption (cumulative)




(b) Investment (cumulative)

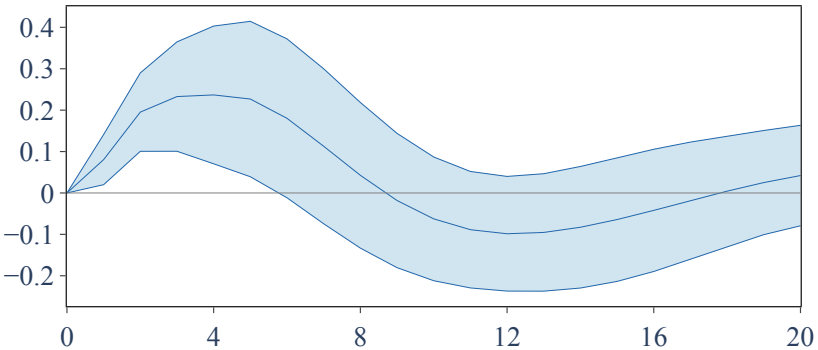



(c) **Output growth** (cumulative)

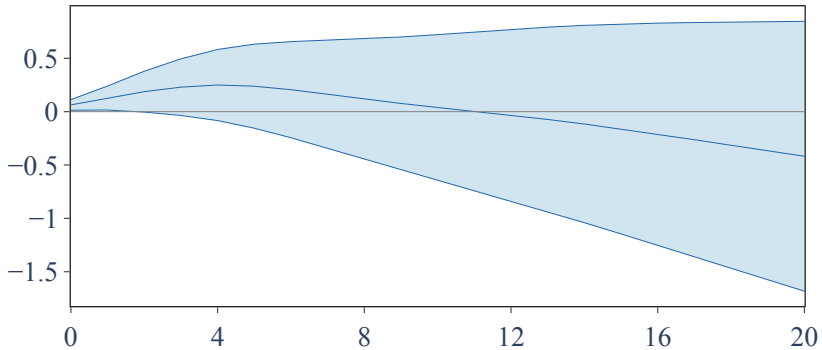



(d) **AI Economy Score**

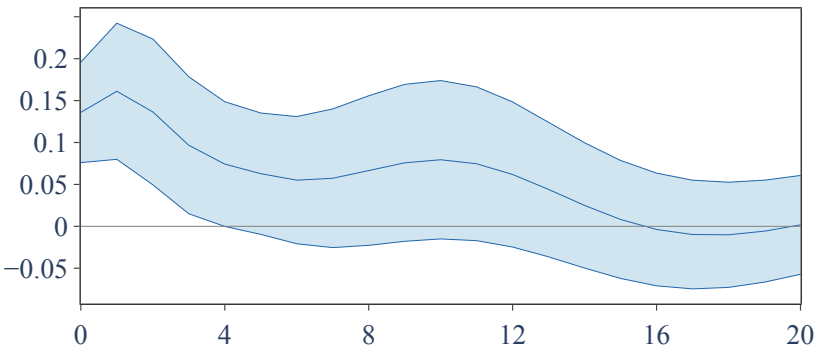
# VAR Analysis of Macroeconomic Implications of Changes in AI Economy Score 2/2



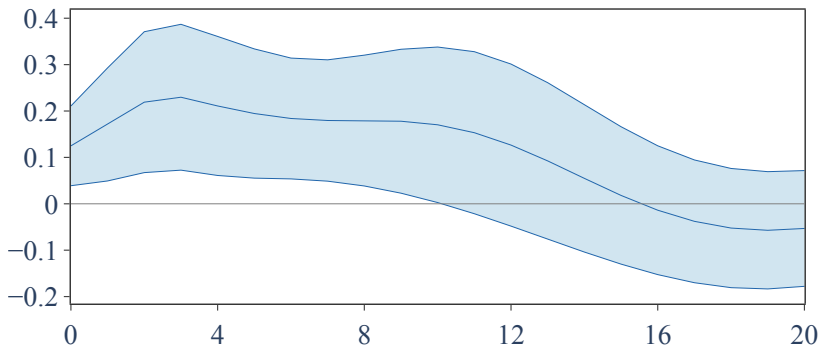

(e) Prices




(f) Excess market return (cumulative)




(g) Ten-year treasury yield




(h) Federal funds rate

# Behavioral Biases: Optimists, Realists, and Pessimists

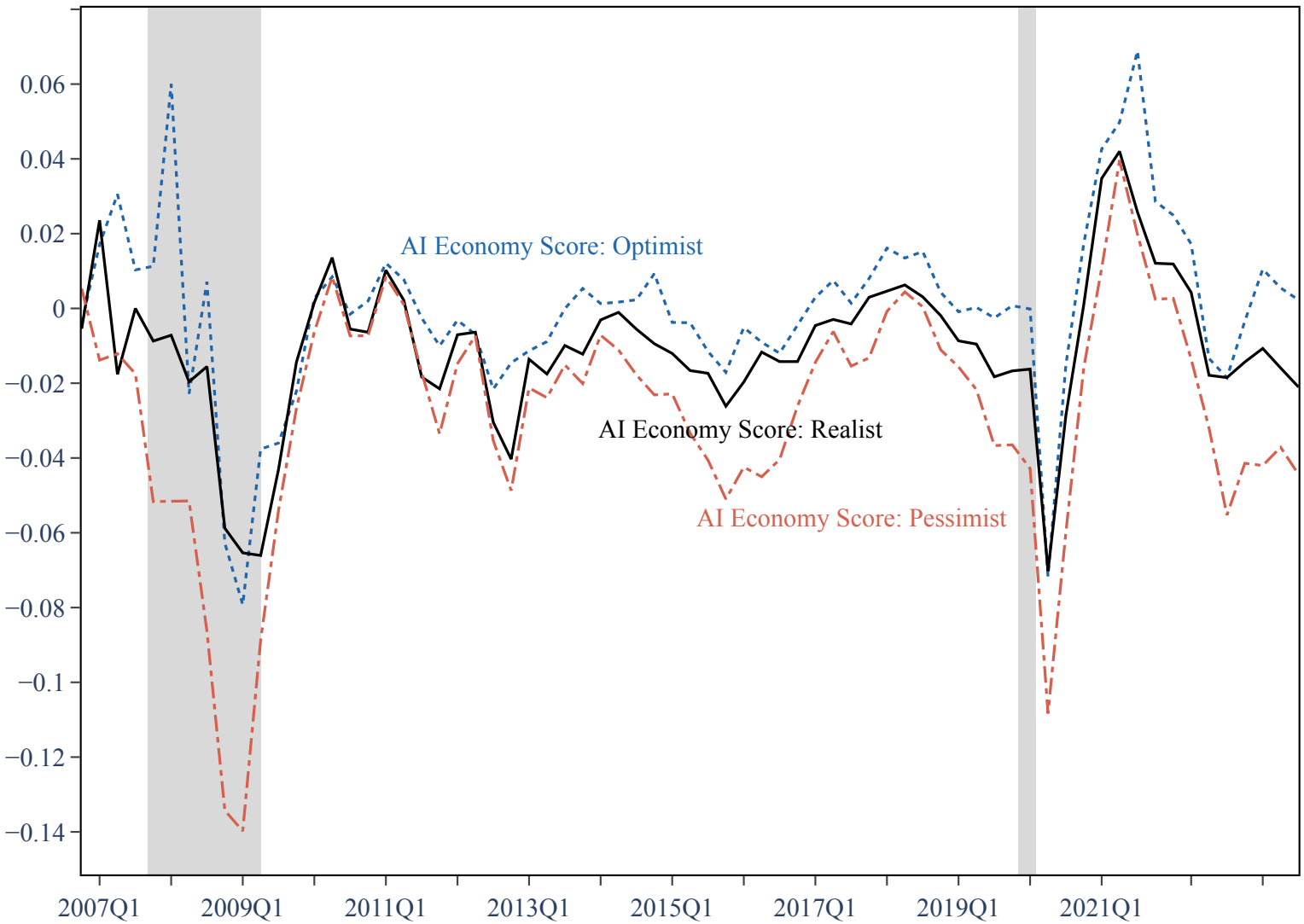
# Classifying Managers as Optimists, Realists, and Pessimists

- We explore whether the predictions are driven by beliefs/noise or by news/information – related to an important debate in macroeconomics (e.g., Barsky and Sims, 2011; Chahrour and Jurado, 2018; Schmitt-Grohé and Uribe, 2012; Blanchard et. al, 2013)

- Based on their prediction bias over the past eight quarters, we classify firms into four time-varying quartiles: Optimists (top quartile), Pessimists (bottom quartile), and Realists (the two middle quartiles).

- We then construct the corresponding AI Economy Scores for each group:
  - AI Economy Score – Optimists
  - AI Economy Score – Pessimists
  - AI Economy Score – Realists

# AI Economy Scores from different groups of firms



# Performance of the Three AI Economy Scores

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | | *Real GDP:* Next Quarter | | | |
| *AI Economy Score* | | 0.341*** | | | | |
| | | (3.15) | | | | |
| *AI Economy Score: Optimist* | | | 0.180** | | | 0.111 |
| | | | (2.62) | | | (1.05) |
| *AI Economy Score: Pessimist* | | | | 0.186** | | 0.106 |
| | | | | (2.52) | | (1.40) |
| *AI Economy Score: Realist* | | | | | 0.266** | 0.0528 |
| | | | | | (2.64) | (0.31) |
| R-squared | 0.455 | 0.494 | 0.487 | 0.483 | 0.480 | 0.495 |
| Observations | 69 | 69 | 69 | 69 | 69 | 69 |
| | | Controlled for term Spread, Real FFR, GZ Spread. | | | | |

- O, P, R can all predict (but not better than collective)
- Relative percentage contributions to R-squared are **37.6% (O), 32.9% (P), and 29.4% (R)**

# Multi-dimensional Information

# Integrating Multi-dimensional Information

- Managerial discussions provide insights into firms' performance, operations, policies, industry trends, and economic conditions.
- **Implicit information about future economic growth can be extracted beyond explicit** managerial forecasts.
- ChatGPT is prompted with **14 questions** covering expectations regarding:
  - U.S. and global economic optimism
  - Firm and industry financial prospects
  - Earnings, revenue, investments
  - Wages, salaries, and number of employees
  - Demand, production quantity, and pricing
  - Input costs and cost of capital

# AI Weighted Score Construction

- We create a new measure, *AI Weighted Score*, as a **linear combination of 14 AI scores**.
- The measure is based on the following panel regression:

$$Sales_{i,s} = \sum_{k=1}^{14} \beta_{k,t} AI\ Score_{k,i,s-1} + \varepsilon_{i,s}, \quad s \leq t-1.$$

- This regression forecasts firm sales **using the previous quarter's AI scores.**
- Coefficients ($\beta_{k,t}$) are estimated without look-ahead bias.

# AI Weighted Score Application

- The weighted score is computed at both national and industry levels:

$$\textit{AI Weighted Score}_t = \sum_{k=1}^{14} \beta_{k,t} \textit{AI Score}_{k,t}, \tag{1}$$

$$\textit{AI Ind. Weighted Score}_{j,t} = \sum_{k=1}^{14} \beta_{k,t} \textit{AI Ind. Score}_{k,j,t}. \tag{2}$$

- These scores help predict economic growth at national and industry levels.
- The *AI Weighted Score* **outperforms the AI Economy Score** in predicting future GDP growth.

# Weighted Score Performs Better

Panel A: National Level

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | | | *Real GDP* | | |
| | 1 qtr | 2 qtrs | 3 qtrs | 4 qtrs | 12 qtrs | 16 qtrs |
| *Term Spread* | 0.172 | -0.108 | -0.152 | -0.287 | -0.994*** | -0.399 |
| | (0.40) | (-0.39) | (-0.43) | (-0.85) | (-5.43) | (-0.81) |
| *Real FFR* | -0.0480 | -0.283* | -0.469*** | -0.738*** | -1.99*** | -0.0175*** |
| | (-0.22) | (-1.86) | (-2.75) | (-4.01) | (-14.38) | (-5.98) |
| *GZ Spread* | -0.891*** | -0.722*** | -0.867*** | -0.935*** | -0.568*** | -1.09*** |
| | (-3.51) | (-2.97) | (-3.70) | (-3.12) | (-3.05) | (-3.97) |
| *AI Weighted Score* | 0.185*** | 0.297** | 0.306*** | 0.263*** | 0.183*** | 0.113 |
| | (3.51) | (2.45) | (3.02) | (2.71) | (2.94) | (1.04) |
| R-squared | 0.485 | 0.520 | 0.580 | 0.587 | 0.795 | 0.741 |
| Observations | 70 | 69 | 68 | 67 | 59 | 55 |

Incorporating **broader range of information can enhance predictive power** for GDP growth, especially for longer-term.

# Robustness

# Robustness Tests

1. Does the *AI Economy Score* provide incremental information compared to Survey Forecasts?
2. Does it work in a subsample with dates, person, product, and firm identity masked?
3. Could researchers use other large language models?
4. Address temporal data leakage - using smaller llm models (work in progress)

We encourage **inputs from audience** on more robustness tests.

# Horse Race with Survey

| | (1) | (2) | (3) |
|---|---|---|---|
| | *Real GDP:* Next Quarter | | |
| *Term Spread* | -0.284 | -0.376 | -0.287 |
| | (-0.85) | (-1.00) | (-0.81) |
| *Real FFR* | -0.157 | -0.163 | -0.157 |
| | (-0.82) | (-0.74) | (-0.82) |
| *AI Economy Score* | 0.625*** | | 0.609*** |
| | (6.83) | | (2.91) |
| *SPF-Forecasted rGDP* | | 0.680*** | 0.0272 |
| | | (3.14) | (0.12) |
| R-squared | 0.448 | 0.338 | 0.448 |
| Observations | 72 | 72 | 72 |

*AI Economy Score* is a **stronger predictor of real GDP** for the near future than survey-based forecasts.

# Masked-Identity Tests

| | (1) | (2) | (3) |
|---|---|---|---|
| | *Real GDP:* Next Quarter | | |
| *Term Spread* | 0.371 | -0.153 | 0.235 |
| | (0.82) | (-0.42) | (0.59) |
| *Real FFR* | 0.0493 | -0.00580 | 0.0594 |
| | (0.21) | (-0.03) | (0.26) |
| *GZ Spread* | -1.320*** | | -0.891*** |
| | (-4.48) | | (-3.74) |
| *AI Economy Score_masked* | | 0.762*** | 0.356** |
| | | (4.33) | (2.21) |
| R-squared | 0.454 | 0.413 | 0.484 |
| Observations | 72 | 72 | 72 |

We mask **person, firm, product, and date** information.
Alleviates concerns about look-ahead bias.

# Another LLM: Llama3

Llama 3 (-8b) is the leading open-source LLM.

| | (1) | (2) | (3) |
|---|---|---|---|
| | *Real GDP:* Next Quarter | | |
| *Term Spread* | 0.371 | -0.414 | 0.306 |
| | (0.82) | (-1.05) | (0.70) |
| *Real FFR* | 0.0493 | -0.0627 | 0.130 |
| | (0.21) | (-0.26) | (0.56) |
| *GZ Spread* | -1.320*** | | -1.233*** |
| | (-4.48) | | (-4.61) |
| *AI Economy Score_Llama* | | 1.023*** | 0.714*** |
| | | (3.46) | (3.39) |
| R-squared | 0.454 | 0.207 | 0.502 |
| Observations | 72 | 72 | 72 |

Predictive ability for is slightly weaker than that of the score generated by ChatGPT.

# Conclusion

- In this paper, we propose **AI Economy Score**, a novel measure capturing the average managerial expectation for the US economy in the next quarter.

- Our findings demonstrate that the AI Economy Score is a strong predictor of future economic activities, including GDP growth, production, employment, and wages, providing additional predictive power beyond those of existing benchmark measures.

- The information uncovered from managerial expectations on the future economic environment can be instrumental for decision making at both the macro and micro levels.

Thank you!