Harnessing Generative AI for Economic Insights*

Manish Jha†
Georgia State

Jialin Qian ‡
Rochester Inst. of Tech.

Michael Weber§
Purdue and NBER

Baozhong Yang¶
Georgia State

September 2025

Abstract

We use generative AI to extract managerial expectations about their economic outlook from over 120,000 corporate conference call transcripts. The overall measure, *AI Economy Score*, robustly predicts future economic indicators such as GDP growth, production, and employment, both in the short term and to 10 quarters. This predictive power is incremental to that of existing measures, including survey forecasts. Moreover, industry and firm-level measures provide valuable information about sector-specific and individual firm activities. A composite measure that integrates managerial expectations about firm, industry, and macroeconomic conditions further significantly improves the forecasting power and predictive horizon of national and sectoral GDP growth. Our findings suggest that managerial expectations carry unique insights about economic activities, with implications for both macroeconomic and microeconomic decision-making.

JEL Classification: C81, D2, E2, E3, E44

Keywords: Artificial Intelligence, ChatGPT, Generative AI, Large Language Model, Macroeconomic Forecast, Managerial Expectations, Textual analysis

*We are grateful to Jiayin Hu, Sophia Kazinnik (discussant), Sophia Zhengzi Li, Dexin Zhou (discussant), and participants at the NBER Summer Institute, China International Conference in Finance, Fidelity Quantitative Research, Financial Intermediation Research Society Conference, Georgia State University, Indian School of Business Summer Research Conference, International Conference on Information Technology and Quantitative Management, McMaster University, Peking University, Rutgers University, Society of Quantitative Analysts Conference, University of Oxford, and Winter Institute of AISSEAI. All errors are our own.

†J. Mack Robinson College of Business, Georgia State University, Email: mjha@gsu.edu

‡Saunders College of Business, Rochester Institute of Technology, Email: jqian@saunders.rit.edu

§Purdue University Daniels School of Business and NBER, Email: weber366@purdue.edu

¶J. Mack Robinson College of Business, Georgia State University, Email: bzyang@gsu.edu

Acknowledgements and Disclosure of Potential Conflicts of Interest

The authors did not receive any financial support for this research. The authors declare that they have no relevant financial relationships or other potential conflicts of interest to disclose.

1. Introduction

Managers of publicly-listed corporations have a large impact on the economy because their corporations employ almost 30% of total non-farm employment in the U.S and are responsible for more than one quarter of GDP (Schlingemann and Stulz, 2022). For this reason, the government and public often listens carefully to their opinions and the media features them prominently.[1] It is well-documented that the expectations of households have important influences on future economic activities because consumption and other activities depend on the information and beliefs contained in their expectations (e.g., Barsky and Sims, 2012; Chahrour and Jurado, 2018; Weber et al., 2022). By the same principle, managerial expectations should also contain valuable information for the economy at both the macro and micro level given their relevance for aggregate employment but also the fact that corporate investment is the most volatile component of aggregate GDP. However, surveys of managers are costly and challenging to conduct on a large scale, a constraint for their wide application.[2] In this paper, we apply the newly available generative AI tools to obtain managerial expectations on a wide range of economic factors from conference call transcripts. Relying on publicly available data, our approach provides an easily accessible set of expectation variables to researchers and policy makers.

We construct managerial expectations building on the approach in Jha, Qian, Weber, and Yang (2023). In particular, we feed corporate conference call transcripts to a generative AI model and prompt it to provide answers regarding managerial expectations of future states of the economy and the firm. This approach generates a panel of answers regarding Gross Domestic

[1] For example, the Wall Street Journal highlighted JPMorgan Chase Chief Executive Jamie Dimon's perspective on U.S. interest rates (Saeedy, 2024). Additionally, CEOs of major companies such as Blackstone, Dell, Disney, Pepsi, and Tesla frequently serve on the President's economic advisory council (Gernon, 2017).

[2] Some examples include the CFO survey, conducted by Duke University's Fuqua School of Business and the Federal Reserve Banks for Richmond and Atlanta: https://cfosurvey.fuqua.duke.edu/; and the Survey of Firms' Inflation Expectations (SoFIE), conducted by Federal Researve Bank of Cleveland: https://www.clevelandfed.org/indicators-and-data/survey-of-firms-inflation-expectations. Important contributions on how firms form their expectations are Graham and Harvey (2001); Ben-David, Graham, and Harvey (2013); Coibion, Gorodnichenko, and Kumar (2018); Coibion, Gorodnichenko, and Ropele (2020). Candia, Coibion, and Gorodnichenko (2023) provide an excellent recent overview of this literature.

Product (GDP) growth, production, employment, and investment, from more than 120,000 conference calls, representing 5,513 unique companies. We then aggregate these answers at the national and industry-sector levels to obtain economic measures that can be indicative of economic activities at these levels.

We first focus on the main measure, the *AI Economy Score*, which captures the average managerial expectation for the US economy in the next quarter. The *AI Economy Score* is strongly predictive of next quarter's real GDP, adding an R-squared of 4% to a model with the common predictors including the *Term Spread*, *Real FFR*, *GZ Spread* (Gilchrist and Zakrajšek, 2012), and lagged GDPs. This predictive power is also moderately persistent, lasting for 4 quarters. The *AI Economy Score* is also significantly associated with other future aggregate economic variables, including industrial production, employment, and wages, for up to 10 quarters in the future.

To filter out the influence of other current economic factors, we employ a vector autoregression (VAR) framework to study the impulse responses to changes in the managerial expectation score. We find that shocks to the *AI Economy Score* orthogonal to current economic activities significantly influence future activities. A positive shock to the expectation score predicts higher consumption, investment, and output growth for 8 or more quarters. It is also positively associated with higher inflation, excess market returns, and long-term government bond yields. The increased economic activity induced by the shock also also trigger tighter monetary policy.

Our approach also allows the construction of more micro-level expectation variables, such as an AI-based industry-level economy score, by aggregating firm-level measures to the sector level. Such scores exhibit substantial cross-sectional heterogeneity, with both pessimistic and optimistic outlooks for different industries at the same point in time. The aggregate and industry economy scores possess independent predictive power of future economic activities across industry sectors. In fact, the predictive power of the industry-level scores for future GDP growth lasts 4 years, suggesting that industry-level expectations have the potential to aid long-term economic decisions.

We can also take one step further and zoom in at the most micro-level, i.e., consider the firm-level managerial expectations. We consider AI-based expectation scores in predicting the top- and bottom-line numbers, i.e., sales and earnings, for firms. We find that both industry-level and firm-level scores forecast future firm-level activities, with firm-level expectations carrying predictive power for up to 4 years.

Our methods also enable the direct construction of managerial expectations of a host of economically related variables, such as employment, wages, industrial production, and others. We verify that these expectation scores provide significant forecasting power for the corresponding future economic variables at both the aggregate and industry levels, and in both the short and long term. While these expectation scores are correlated with the AI economy scores, these variables contain additional informatoin. For example, AI employment scores are more powerful predictors of future employment and unemployment rates than the overall economy score.

Furthermore, beyond explicit forecasts, managerial expectations about other economic variables also contain implicit signals about future economic growth. To extract this information, we prompt ChatGPT with 14 different questions covering various aspects of managers' expectations, from firm financial prospects to macroeconomic trends. Firm-level AI scores derived from ChatGPT responses are aggregated into industry and national-level scores. A new measure, the *AI Weighted Score*, is constructed using the weights from a firm-quarter-level regression model that relates these scores to future firm sales. Results show that this score significantly improves predictions of GDP growth at both national and industry levels, outperforming simpler AI-based measures. In particular, the new measure can forecast national GDP growth for up to 14 quarters, significantly expanding the horizon relative to the original *AI Economy Score*. The most influential scores within the weighted score include revenue, production, wages, employment, industry prospects, and capital expenditures. These findings suggest that leveraging AI to extract multidimensional managerial expectations from earnings calls can enhance economic forecasting, particularly for long-term growth projections.

Why are managerial expectations so powerful in forecasting future economic activities? One important reason is that managers can act upon their expectations or beliefs and their actions carry significant economic impacts. For example, Jha, Qian, Weber, and Yang (2023) show that managerial expectations of corporate investment policies significantly affect long-term investment activities. Another potential reason is that managers are privy to first-hand, detailed economic data and information that may not be available to the public. Therefore, aggregating such information and insights from managers would be particularly fruitful. Overall, our study shows that the new AI tools allow easy extraction of managerial expectations on a whole range of economic indicators, providing a new set of forward-looking information to researchers, investors, and regulators. While such information is mostly positive in nature, given its significant predictive power, it can also be instrumental for normative policies.

Our paper contributes to the extensive literature on economic expectations and real economic activities. Related studies have examined how news about future economic conditions impacts business cycles and macroeconomic stability (Barsky and Sims, 2011; Chahrour and Jurado, 2018; Schmitt-Grohé and Uribe, 2012; Blanchard, L'Huillier, and Lorenzoni, 2013). Other research has explored how changes in consumer confidence and firm expectations influence economic activities (Barsky and Sims, 2012; Coibion, Gorodnichenko, and Kumar, 2018; Coibion, Gorodnichenko, and Weber, 2022), and how survey data on subjective beliefs affect the business cycle and determinants of inflation expectations (Bhandari, Borovička, and Ho, 2022; Coibion, Gorodnichenko, and Ropele, 2020; Weber, D'Acunto, Gorodnichenko, and Coibion, 2022). Our managerial expectation measures are strongly predictive of future economic activities. They can serve as complements to the standard consumer surveys and surveys of professional forecasters.

A related literature provides variables indicative of economic activities, e.g., Gilchrist and Zakrajšek (2012). Bybee (2023) is a closely related contemporaneous paper. The paper generates survey results from Wall Street Journal news articles. We regard the two papers as complementary since we focus on the expectations retrieved from corporate managers whereas the focus in

Bybee (2023) is on news reports. Our approach generates variables with long-run predictability and also has the capability of generating aggregate, industry-level, state-level, and firm-level expectations, which can be fruitfully employed both at the macro- but also at the micro-level. Additionally, our use of conference call transcripts to gauge national sentiments sets this work apart from other works that rely on book corpora (Jha, Liu, and Manela, 2022) or newspapers (van Binsbergen et al., 2024).

Finally, our paper expands the use of generative AI tools in economics.[3] Korinek (2023) provides a survey of potential uses of generative AI in economic research. Our study and method may help to open more avenues for economic research with the help of AI. For example, a recent paper by Sheng, Sun, Yang, and Zhang (2024) has built on our approach to investigate the reliance on generative AI by investment companies.

2. Data

In this section, we discuss the different datasets we use as well as the variable construction.

2.1. Conference Call Transcripts

We obtain earnings call transcripts for publicly listed firms spanning the years 2006 to 2020 from the SeekingAlpha website and for the period from 2021 to 2023 from the FinancialModeling-Prep website. These transcripts are derived from quarterly earnings call conferences, which are critical events in which firms engage with their investors and analysts following the announcement of their quarterly financial performance. During these calls, managers typically begin with a detailed speech that summarizes the firm's recent performance, discusses their outlook, and addresses any potential concerns regarding the firm's future. After the initial presentation, the floor is opened to analysts and investors, who then ask questions. These questions often pertain

[3]This is a fast-growing literature. A partial list includes: Lopez-Lira and Tang (2023); Hansen and Kazinnik (2023); Kim, Muhn, and Nikolaev (2023, 2024); Jha, Qian, Weber, and Yang (2023); Yang (2023); Li, Mai, Shen, Yang, and Zhang (2023); Ouyang, Yun, and Zheng (2024).

not only to the specifics of the firm's operations but also touch on broader economic conditions. Managers respond to these questions in real time, providing their insights and opinions. This interaction offers a valuable window into the perspectives of top management on both their company's prospects and the general economic environment.

2.2. Real Outcomes

We collect national-level data on GDP, industrial production, employment, federal funds rates, and treasury yields from the Federal Reserve Economic Data (FRED) database managed by the Federal Reserve Bank of St. Louis. We source industrial-level economic indicators from the U.S. Bureau of Economic Analysis (BEA), a division of the U.S. Department of Commerce. Additionally, we gather firm-level financial outcomes from Compustat and the Center for Research in Security Prices (CRSP).

2.3. Survey Data

We use forecasts on economic indicators from the Survey of Professional Forecasters website (SPF) conducted by the Federal Reserve Bank of Philadelphia. Established in 1968, the SPF is one of the oldest quarterly surveys of its kind in the United States. The collected data provides valuable insights into the collective expectations of professional forecasters, helping policymakers, researchers, and the public understand economic trends and make informed decisions. In Section 7.1, we conduct a horserace analysis, comparing the professionals' forecasts with those made by firm managers as retrieved by generative AI.

2.4. Economic Target Variables

Our final sample comprises 149,720 earnings call transcripts spanning the years 2006 to 2023. We aggregate managerial forecasts to the national and industrial level and examine the predictive ability of these measures on future economic development across different horizons. In all levels of analyses, we focus on forecasting the following common economic indicators:

GDP, production, employment, and wages. For firm-level predictions, we use *Value-Add*, or the difference between sales and cost of goods sold, as a proxy for the overall value created by firms and *Sales* as a proxy for production.

We define the realized economic indicators as the logarithm of economic indicators for the next period i we forecast $(t+i)$ over the most recent period with known predictors $(t-1)$ in our regression analyses. We control for common economic predictors in line with Gilchrist and Zakrajšek (2012): *Term Spread*, *Real FFR* and *GZ Spread*. *Term Spread* is the difference between ten-year and three-month constant-maturity Treasury yield. *Real FFR* is the real federal funds rate, earlier works including Bernanke and Blinder (1992) show that the federal funds interest rate offers insights into upcoming shifts in real macroeconomic factors. Additionally, the *GZ Spread* serves as the indicator of movements incredit spreads.

3. Empirical Methodology

In this section, we describe the construction of our key AI-based economic forecast measure, provide validations of the measure, and present summary statistics.

3.1. AI Economy Score

To infer economic insights from texts, we use the leading generative AI model, ChatGPT, developed by OpenAI, based on their Generative Pre-trained Transformer (GPT) model series. GPT's transformer architecture, utilizing layers of self-attention mechanisms, excels at understanding sophisticated semantics of long texts. Google's BERT, launched in 2018, was the first successful transformer-based natural language processing model (Devlin et al., 2019). Another milestone in transformer-based models is OpenAI's GPT-3, with 175 billion parameters, released in June 2020. ChatGPT was based on the GPT models, further finetuned by reinforcement learning with human feedback. After its release on November 30, 2022, ChatGPT immediately demonstrated its ability to generate detailed responses across various knowledge domains. ChatGPT is particularly

effective at analyzing conference call texts due to its consistency, objectivity, and ability to handle lengthy, complex documents that challenge human comprehension.

We use ChatGPT 3.5 Turbo to process texts, accommodating up to 4,096 tokens, which is roughly equivalent to 3,000 words.[4] For practical purposes, we segment each conference call into parts no longer than 2,500 words, ensuring sufficient space for responses. A typical earnings call, usually around 7,500 words, is divided into three sections. To extract firm's outlook towards the US economy from transcripts, we employ specific prompts customized for ChatGPT as outlined below.

> 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.]*

Based on the model's response in "choice", we extract a choice for each section and assign scores of -1, -0.5, 0, 0.5, and 1, corresponding to the options: Decrease substantially, Decrease, No change, Increase, and Increase substantially. If ChatGPT indicates that "no information is available," we assign a score of zero. We calculate the *AI Economy Score* as the average of these scores from all parts of a single earnings call, creating a metric at the firm-quarter level. We find our main findings to be consistent across different methods of aggregating scores from text

[4] Due to the size of the corpus of conference call transcripts, using the more advanced GPT-4 models would be too costly for our study.

sections.

3.2. Validation

Validating machine learning outputs is crucial for ensuring accuracy and reliability. We employ several methods to validate the responses generated by ChatGPT: (i) Manual review and frequency analysis: We manually read through a random sample of around a thousand conference calls and ChatGPT responses to verify the accuracy and relevance of the generated content. We also examine the most commonly used words associated with both low and high AI Economy Scores to gauge the consistency and appropriateness of ChatGPT's outputs; (ii) Trend analysis: We analyze trends over time and across various industries to assess the coherence and relevance of ChatGPT's responses within different contexts.

3.2.1. Manual Review and Frequency Analysis. We first manually go over a random sample of the conference calls and ChatGPT responses, verifying the accuracy and relevance of the generated content. We find ChatGPT to consistently generate high-quality responses similar to human assessments. Appendix B provides example texts associated with high and low AI scores respectively.

In our prompt to ChatGPT, we ask for responses in the format of "choice – explanation." We analyze these explanations using two approaches. First, as shown in Figure 1, we categorize all explanations associated with positive or negative AI Economy Scores into distinct topics.

[Insert Figure 1 Here]

Second, we focus on the explanations linked to the responses "significantly decrease" and "decrease" (and separately for "significantly increase" and "increase"). From this set of explanation texts, we remove stopwords and non-English words using the dictionaries available in the Python Natural Language Toolkit package. We remove proper nouns and then lemmatize each word to account for differing grammatical noun and verb forms. Finally, we use the CountVectorizer

function in the Python package sklearn to create the frequency of the most common 3-grams and 4-grams for firms with low and high AI economy scores, respectively.

[Insert Table 1 Here]

Table 1 shows the top 10 phrases (3- and 4-grams) associated with low and high AI Economy scores. As the table shows, the n-grams reflect relevant themes. For low AI Economy scores, the n-grams indicate a challenging economic environment, adverse market conditions, and a tough business climate. The conference calls in these cases often discuss global economic conditions, firms being cautious, currency exchange risks, and decreased optimism. In contrast, the n-grams associated with high AI Economy scores reflect optimism, strong revenue growth, and positive financial performance. Discussions in these calls highlight sales growth, improved market conditions, organic revenue growth, and overall positive business performance.

3.2.2. Trend Analysis. We next examine the time trends in AI economy scores and how they match with actual GDP growth, and the predicted GDP growth based on the Survey of Professional Forecasters (SPF). In Panel (a) of Figure 2, we present a comparison between the *AI Economy Score* and the actual GDP growth for the subsequent quarter throughout the entire sample period. The two lines generally exhibit strong alignment, indicating that the *AI Economy Score* effectively forecasts the fluctuations in GDP growth, particularly during the 2008 financial crisis. Although the score does not immediately predict GDP growth for the first quarter of 2020, this discrepancy is understandable given the unforeseen impact of the COVID-19 pandemic. Overall, the evidence suggests that the AI economy score is a strong predictor of future economic conditions in the US.

Panel (b) illustrates the time-series trends of the *AI Economy Score* alongside the SPF-predicted real GDP growth for the same quarter, over the entire sample period. These two lines also demonstrate strong comovement, including during the COVID-19 downturn. Since the AI economy score aggregates opinions of corporate managers, its high correlation with the SPF

forecasts suggests that managers and professional forecasters indeed have similar outlooks on the economy. However, we note that the AI forecasts likely carry additional information beyond those in the SPF forecasts given the differences in the information sets and skillsets of managers and forecasters (Section 7.1).

[Insert Figure 2 Here]

Next, we analyze the heterogeneity in the *AI Economy Score* across industries. Figure 3 illustrates the changes in the *AI Economy Score* from 2006 to 2023 for 19 industries classified under the North American Industry Classification System (NAICS). Notably, during the recessions of 2007-2009 and 2020, a significant increase in negative AI Economy Scores arises, depicted in red, across all industries. This trend highlights the widespread economic downturns experienced during these periods.

[Insert Figure 3 Here]

Substantial heterogeneity for AI forecasts across industries is present. For instance, during the 2007-2009 recession, Retail, Transportation and Warehousing, and Educational Services sectors were most severely affected. In contrast, Healthcare and Utilities industries demonstrated relative resilience. Similarly, in 2020, the Retail sector faced the most substantial negative impact, whereas Healthcare services maintained a relatively positive outlook on the economy. Post-recession recovery is evident in the positive *AI Economy Score*s, shown in blue. After the COVID-19 pandemic in 2021, Technical Services exhibited the most optimistic recovery. In the years following the 2007-2009 recession, Technical Services and Manufacturing sectors showed the highest levels of optimism. These findings align well with prior expectations and actual historical events.

3.3. Summary Statistics

Table 2 displays the descriptive statistics of the *AI Economy Score*, the realized values of various macroeconomic indicators, and predictor variables of future economic conditions at the

aggregate level. Our sample period is from 2006 to 2023.

[Insert Table 2 Here]

Over the 72 quarters in our same period, the average *AI Economy Score* is -0.0131, with a standard deviation of 0.022. The distribution of AI-predicted managerial forecasts is less skewed than that of survey forecasts. In our predictive regressions, we use the logarithm growth rates of the economic indicators (including *Real GDP*, *Industrial Production*, *Payroll Employment*, and *Wages*) for the next period over the most recently known period. *Term Spread* and *Real FFR* are quarterly measures obtained from averaging monthly values. The credit spread index *GZ Spread* is the quarterly GZ spread obtained by taking the average of the monthly values.

4. AI Forecasts on Economic Indicators

4.1. AI-Predicted Economy Score and Realized Real GDP

To evaluate the predictive ability of the AI-Predicted Economy Score for economic activities, we estimate the following forecasting specification:

$$ln\frac{Real\ GDP_{t+1}}{Real\ GDP_{t-1}} = \alpha + \beta_1 AI\ Economy\ Score_t + \beta_2 Term\ Spread_t + \beta_3 Real\ FFR_t + \beta_4 GZ\ Spread_t + \sum_{i=1}^{4} ln\frac{Real\ GDP_{t-i}}{Real\ GDP_{t-i-1}} + \varepsilon_t \quad (1)$$

where *Term Spread* is the difference between ten-year and three-month constant-maturity Treasury yield, *Real FFR* denotes real fed funds rate, and *GZ Spread* is the credit spread index. Four lags of *Real GDP* growth are included as controls.

Table 3 details the comparison of the predictive power of two financial indicators, the *GZ Spread* and the managerial economic forecast, *AI Economy Score*, for real GDP growth. We focus on the one-quarter and longer-term (up to 6 quarters ahead) forecast horizons and report the

coefficients associated with the financial indicators, as well as the goodness of fit as measured by R^2.

Panel A in Table 3 presents the results for the one-quarter ahead forecasting period. Column (1) shows that the *GZ Spread* has strong explanatory power for real GDP, consistent with Gilchrist and Zakrajšek (2012), whereas the *Term Spread* and *Real FFR* do not exhibit short-run forecasting ability during our sample period. In column (2), we replace the *GZ Spread* with the *AI Economy Score*. The coefficient on the *AI Economy Score* is positive and significant at the 1% level, indicating that our *AI Economy Score* is a strong predictor of the next quarter's real GDP growth. The effect is also economically significant - one standard deviation increase in the *AI Economy Score* implies a 1.38 percentage point (0.625*0.022) increase in the growth rate of real GDP over the subsequent quarter. In column (3), we include both the *GZ Spread* and the *AI Economy Score* in the predictor set. The *AI Economy Score* remains statistically significant at the 1% level. The R^2 increases from 0.454 in column (1) to 0.493 in column (3), indicating that the *AI Economy Score* adds substantial incremental forecasting power to those of existing predictors.

In Panel B, we examine the long-term predictive power of the *AI Economy Score*. Our analysis reveals that the *AI Economy Score* retains its predictive power regarding real GDP growth for up to four quarters. In contrast, the *GZ Spread* remains a highly significant predictor of future real GDP for up to six quarters. Although the *GZ Spread* demonstrates stronger predictive power at longer horizons, our results indicate that the *AI Economy Score* provides valuable information about long-term economic prospects that extends beyond the scope of existing predictors.

The results in Table 3 show that the *AI Economy Score* contains additional information about the future economy and provides extra predictive power for real GDP, beyond the common predictors documented in the literature (Bernanke and Blinder, 1992; Gilchrist and Zakrajšek, 2012).

4.2. AI-Predicted Economy Score and Various Economic Indicators

As managers usually touch on multiple topics related to the economy during earnings call conferences, the economic forecasting score extracted from these calls can contain forecasts on different aspects of economic activities. In Table 4, we test the forecasting ability of the *AI Economy Score* on Industrial Production, Employment, and Wages over the short and long term with the same specifications as in Table 3. In columns (1), (4), and (7) of Panel A, we include the three existing economic predictors: *Term Spread*, *Real FFR*, and *GZ Spread*. The results show that the *GZ Spread* is a strong predictor of all three economic indicators for the next quarter, whereas the *Term Spread* and *Real FFR* can only predict Industrial Production for the next quarter. We replace the *GZ Spread* with *AI Economy Score* in columns (2), (5), and (8) to compare the predictive power of the two variables. The goodness-of-fit comparison implies that the *AI Economy Score* exhibits higher predictive power for Employment and Wages, whereas the *GZ Spread* performs better for Industrial Production. For example, the inclusion of the *AI Economy Score* to predict employment yields an R^2 of 36.1%, a 25% improvement compared to that of the *GZ Spread*. Also, simultaneously adding both the *GZ Spread* and *AI Economy Score* in the model reduces the predictive power of the *GZ Spread*, as shown in columns (3), (6), and (9). The magnitude of the coefficients attached to the *AI Economy Score* is also economically significant: a one-standard-deviation increase of *AI Economy Score* is associated with 2.93%, 1.41%, and 0.24% increases in the growth rate of industrial Production, Employment, and Wages over the next quarter, which corresponds to 86%, 64%, and 48% of a standard deviation in the growth rate of Industrial Production, Employment, and Wages.

[Insert Table 4 Here]

In Panel B, we study the longer term predictive power of the *AI Economy Score* on industrial production, employment, and wages. The results imply that the *AI Economy Score* can predict industrial production for about seven quarters, although the statistical power weakens for

predictions four to seven quarters ahead. Additionally, the results show that the *AI Economy Score* can predict employment for up to ten quarters and wages for approximately eight quarters. In contrast, while the *GZ Spread* can predict industrial production for six quarters and employment for five to ten quarters ahead, it fails to predict near-term employment and wages for any periods longer than one quarter.

Overall, the results in Table 4 indicate that the *AI Economy Score* is a strong predictor of multiple aspects of economic activities over a long period, including industrial production, employment, and wages, even when controlling for known predictors of economic prospects. Additionally, the forecasting ability of the *AI Economy Score* is complementary to the *GZ Spread* regarding employment and wages. These results are plausible because the AI-predicted score is obtained from managers' outlooks and conversations with investors and analysts, and it should contain more information about firms' labor-related policies as compared to conditions in credit markets.

4.3. VAR Analysis

In this section, we employ a vector autoregression (VAR) framework to analyze the impulse responses to shocks in the managerial expectation scores. We add the *AI Economy Score* to a standard VAR model, which includes the following endogenous variables: (a) the real personal consumption expenditures; (b) the real business fixed investment; (c) the real GDP; (d) inflation, measured by the log-difference of the GDP price deflator; (e) the *AI Economy Score*; (f) the quarterly value-weighted excess stock market return; (g) the nominal ten-year Treasury yield; and (h) the nominal effective federal funds rate. Consumption, fixed investment, and real GDP are measured in log differences. The VAR model is estimated using two lags for each endogenous variable.

Figure 4 presents the impulse response functions of the endogenous variables in response to an orthogonalized shock to managerial expectations, as measured by the *AI Economy Score*. An unexpected one-standard-deviation increase in the *AI Economy Score* significantly boosts

multiple economic indicators, including consumption, investments, and output, over the next 20 quarters. The magnitude of these responses is substantial. Following the unexpected shock to managerial forecasts, consumption and output rise by about 2 percentage points above trend, while investment shows an even stronger response, increasing by approximately 6 percentage points above the trend. This economic surge leads to significant inflation in the subsequent quarters.

After a shock to managerial economic forecasts, monetary policy tightens, reflected by an immediate rise in the federal funds rate that persists for the next 12 quarters. Additionally, the ten-year treasury yield gradually increases over time. The stock market sees a slight uptick during the 8 quarters following the shock. In sum, the VAR results show that the *AI Economy Score* yields independent and substantial impact on a host of economic variables.

[Insert Figure 4 Here]

4.4. AI Forecasts of Additional Macro Variables

We also analyze three other macro variables: (i) aggregate production, (ii) aggregate employment, and (iii) aggregate wages. All three are included in both the Duke CFO Survey and the industry-level data available in the BEA database. To generate AI scores for these variables, we follow a similar process as described in Section 3.1. However, we replace the questions with "Over the next quarter, how does the firm anticipate a change in ..." (i) production quantity of its products?, (ii) number of employees?, and (iii) wages and salaries expenses?. In unreported results, we find that the three AI scores significantly predict real outcomes, even after controlling for *Term Spread*, *Real FFR*, and *GZ Spread*, in a setup similar to Table 3.

Appendix Figure IA.1 presents the time series variations in AI scores and aggregate production, employment, and wages. In the graphs, the AI conference transcript-based scores closely resemble the actual outcomes reported by the BEA. Both the AI-based scores and the actual outcomes show declines during the recessions of 2007-09 and 2020. Again, we see the change

in AI scores are a bit elayed during Covid-19 pandemic, as it would be hard for managers to anticipate. While the production and employment figures are more volatile, wages are sticky and do not dip or increase substantially from one quarter to another.

Appendix Figure IA.2 illustrates the variations in the three AI scores across different industries. Darker red shades indicate periods of contractions, whereas blue shades signify expansions. The impact varies across industries. During the 2007-09 crisis, the three AI scores (production, employment, and wages) dropped significantly in the retail and educational services sectors. After the COVID-19 recovery in 2021, the educational services and information technology sectors saw the largest increases in these AI scores. Notably, the darker red representing aggregate wage expenses for companies in the accommodation and food services industries captures the impact of travel restrictions in 2020.

5. AI Forecasts on Economic Indicators at Micro Levels

Since our national-level *AI Economy Score* is aggregated from firm-level managerial measures, we can obtain predictors at more disaggregated levels, such as the industry-level and firm-level. In this section, we analyze the predictive power of managerial forecasts for the US economy at the micro level.

5.1. AI Industry Score and Industrial Economic Indicators

We classify firms into 19 sectors based on the NAICS industrial classification standard, then convert the firm-level AI-predicted scores to industry-level scores by calculating the equally-weighted average within each sector. We use the AI-predicted score for the US economy as our baseline, similar to the national-level analysis. The results for other scores are displayed in Appendix D. Panel A in Table 5 presents the summary statistics of the main variables used at the industry level. Among the predicted variables, the industrial employment and wage data are only available annually. Therefore, we analyze employment and wage data on an annual basis. We

construct annual AI-predicted scores by averaging the quarterly scores for each year. Similarly, the annual *Term Spread* and *Real FFR* are obtained by averaging the monthly measures for each year. The distribution of AI industry scores is comparable to the national scores.

[Insert Table 5 Here]

Table 6 presents the coefficient estimates of various predictors for real sectoral GDP. Panel A displays the short-term results. The improvement in goodness of fit brought by the *AI Economy Score* matches that of the *GZ Spread*, as shown in columns (1) and (2). Columns (2) and (3) in Panel A indicate that both the aggregate and industry-level AI Economy Scores are strong predictors of industrial real GDP. Specifically, a one-standard-deviation increase in the *AI Economy Score* or the *AI Industry Score* is associated with a 1.48% or 0.9% increase, respectively, in the growth rate of real sector GDP for the next quarter. In Column (4), which incorporates all predictors, the *AI Industry Score* shows additional predictive power for industry real GDP, even when controlling for all known predictors. Panel B of Table 6 reveals that the aggregate *AI Economy Score* can predict future industrial real GDP up to three quarters ahead, while the *AI Industry Score* possesses predictive ability for up to 16 quarters, longer than the forecasting periods of the *GZ Spread*, which is 8 quarters.

[Insert Table 6 Here]

Similarly, Table 7 demonstrates that the industrial-level and aggregate scores can predict other economic indicators at the industry level. For example, column (3) in Panels A and B show that the *AI Industry Score* is positively and significantly associated with employment and wages in the corresponding industries. Compared to the coefficient attached to the *GZ Spread* in column (1), the *AI Economy Score* in column (2) exhibits stronger predictive power regarding employment and wages at the industry level. When including all the predictors in one specification, both *AI Economy Score* and *AI Industry Score* display strong predictive power. The long-term real GDP forecasting results at the industry level are presented in Table IA.1 of the Internet Appendix. The

AI Industry Score is a strong predictor of industrial employment and wages for up to four years, even when controlling for other existing predictors.

[Insert Table 7 Here]

5.2. AI Firm Score and Firm Financial Indicators

To further leverage the micro-level predictive power of our measure, we conduct the analysis at the most granular level, focusing on predicting firm-level outcomes. Panel B in Table 5 presents the summary statistics of firm-level variables. In firm-level analyses, we use the score obtained from the question regarding the manager's outlook for the firm they manage as the main forecasting variable, referred to as the *AI Firm Score*. We also omit the longer-term analysis (four quarters ahead in previous tables) at the firm level, as the firm-level data is only available on an annual basis. Additionally, since the firm-level *Wage* variable (xlr or "Staff Expense - Total") in Compustat has limited availability, the sample size is significantly reduced for the firm-level wage analysis.

Table 8 presents the results for predicting value-added, sales, employment, and wages respectively from Panel A to Panel D. We replace the firm-level *AI Economy Score* with the *AI Firm Score* in all specifications of Table 8 because it is a more direct measure of a manager's opinion about their own firm's prospects. This is in contrast to the measure constructed from questions about the overall US economy, which serves as our main measure for a more macro-level setting. In columns (1) to (4), we include only one predictor at a time while controlling for other national-level and firm-level predictors, allowing us to compare their predictive power. Throughout all five columns of the specifications, we control for *Size*, *Book-to-Market*, and *Tangibility*. Firm-fixed effects are included in all specifications. Note that the number of observations may vary in different panels depending on the availability of target and control variables.

[Insert Table 8 Here]

We find that the three AI-based variables, *AI Economy Score*, *AI Industry Score*, and *AI Firm Score*, all forecast value-added, sales, employment, and wages for the next quarter at the firm level, with a significance at the 1% level. The *AI Firm Score* shows the highest predictive power in terms of the improvement in goodness of fit. The magnitude of the coefficients is also economically significant. For example, a one-standard-deviation increase in the *AI Firm Score* is associated with a 6.6% increase in the growth rate of Value-Add and a 5.2% increase in the growth rate of sales for the next quarter, as well as a 6.3% increase in employment and a 5.4% increase in wages over the next year.

The long-term forecasting results at the firm level are presented in Table IA.2 of the Internet Appendix. The *AI Firm Score* shows a strong and positive correlation with firms' value-added, sales, employment, and wage level for up to four years when controlling for other known macroeconomic indicators and firm covariates.

The firm-level results provide further evidence that the managerial forecasts derived from earnings calls by Generative AI tools contain substantial incremental information about both short-term and long-term firm prospects that cannot be explained by the known predictors of the aggregate economy and firm characteristics.

6. Dissecting Managerial Forecasts: Behavioral Biases and Multi-dimensional Insights

6.1. Heterogeneity in Managerial Behavior Types

Managers differ in their beliefs and styles, which leads to variations in their forecasts for the future economy. We next examine the forecasting performance of different types of managers and compare their forecasts during crisis periods and normal times.

We define managerial style based on the difference between their forecasts and the real GDP tercile over the past 8 quarters. Specifically, we classify managerial forecasts into three terciles based on the sign of the AI Economy Score derived from the firm's earnings call for a

given quarter. We also categorize quarterly real GDP, during our sample 2006-2023, into three terciles. The difference between the forecast tercile and the real GDP tercile represents the bias in managerial forecasts.

Next, we calculate each firm's average bias over the past eight quarters and divide the sample into quartiles. If the average bias falls into the top quartile, the forecast is classified as optimistic; if it falls into the bottom quartile, it is classified as pessimistic; the remaining forecasts are categorized as neutral. Finally, we aggregate the AI Economy Score by taking the average across the different managerial types, resulting in three distinct scores: *AI Economy Score: Optimist*, *AI Economy Score: Pessimist*, and *AI Economy Score: Realist*.

Figure 5 shows the time series trend of the three scores. It illustrates the persistence of managerial behavior, where managers who were optimistic in the past tend to remain optimistic in the future, and those who were pessimistic also tend to maintain a pessimistic outlook in subsequent periods. This provides validation for our method of measuring managerial behavior types.

Table 9 presents the regression results. In Panel A, we find that all three scores exhibit strong predictive power for future GDP, with *AI Economy Score: Realist* showing the smallest contribution towards improving variance explanation (increased R-squared). These findings suggest that managerial beliefs indeed influence the real economy.

Panel B shows the results at the micro level, where the dependent variable is real GDP for each quarter, and the *AI Economy Score* is measured at the firm level. We interact the *AI Economy Score* with indicators of optimism and pessimism, while controlling for firm fixed effects to account for unobserved time-varying firm characteristics. Consistent with the macro-level results, we find significant results for *AI Economy Score: Pessimist* again, but no significant results for *AI Economy Score: Optimist*. Interestingly, the interaction between *AI Economy Score: Pessimist* and the crisis period indicator is also significant, suggesting that pessimistic managers perform better at predicting macro economy during crisis periods.

[Insert Table 9 Here]

6.2. Heterogeneity in Business Diversification

Managers in firms with varying levels of concentration may differ in their ability to forecast national economic conditions. One possibility is that managers in more diversified firms have access to a broader set of information channels, as they must engage with various sectors of the economy. This increased exposure could lead to more accurate economic forecasts. However, prior literature suggests that more diversified firms tend to underperform compared to their more focused counterparts due to higher agency costs and fewer growth opportunities (Lang and Stulz, 1994; Berger and Ofek, 1995). If this holds, managers in diversified firms may generate forecasts that are more biased.

To test these two opposing hypotheses, we investigate the relationship between a firm's level of concentration and the predictive power of future economic conditions. We measure concentration level using two proxies. First, we employ the business operation Herfindahl-Hirschman Index (HHI), calculated as the sum of the squared ratio of each business segment's sales to total sales. *High Business Segment HHI* is defined as an indicator variable that takes the value of one if the Business Segment HHI exceeds the median for all firms within a quarter. Second, we use GEO HHI, which is the sum of the squared ratios of the frequency with which each state is mentioned in a firm's 10-K filing relative to the number of different states mentioned (Garcia and Norli, 2012). *High GEO HHI* is an indicator variable that equals one if the GEO HHI is above the median for all firms in a given quarter.

Table 10 presents the results. In columns (1) to (4), I report the interactions between the AI Economy Score and the diversification measures in firm-level panel regressions. All specifications control for the four lags of the dependent variable, other macroeconomic predictors, and firm fixed effects. The four interaction terms consistently exhibit positive and statistically significant coefficients, indicating that the predictive power of economic forecasts is stronger for more focused firms compared to more diversified firms. These results support the hypothesis

that economic forecasts from more diversified firms are more biased.

[Insert Table 10 Here]

6.3. Integrating Multi-dimensional Information from Managerial Forecasts

Managerial discussions in conference calls contain a rich set of information, including expectations about multiple aspects of firms' performance, operations, and policies, industry trends, and economic conditions. These discussions may contain additional implicit information about future economic growth beyond what is contained in explicit managerial economic forecasts. To capture such information, we also extract managers' forecasts regarding changes in various economic and firm-related aspects for the next quarter. In addition to asking about their expectations for optimism in the U.S. economy, we prompt ChatGPT with thirteen other questions covering managers' expectations regarding changes in optimism about the global economy, their firm's financial prospects, their industry's financial prospects, earnings, revenue, investments, wages and salaries, number of employees, demand for their products or services, production quantity, product or service prices, input or commodity prices, and the cost of capital.[5] Based on the responses from ChatGPT, we first obtain firm-level AI scores for each prompted aspect and then construct industry-level and national-level scores by averaging firm-level scores.

To synthesize the information contained in these scores, we compose a new measure, *AI Weighted Score*, that is a linear combination of the 14 scores as follows. First, for each quarter t, we conduct the following panel regression

$$Sales_{i,s} = \sum_{k=1}^{14} \beta_{k,t} AI\ Score_{k,i,s-1} + \varepsilon_{i,s}, \quad s \leq t-1. \tag{2}$$

where $AI\ Score_{k,i,s-1}$ is the firm-level AI Score for the k-th aspect of firm i in quarter $s-1$. In other words, we try to forecast firm sales via a multilinear regression on the 14 firm-level scores

[5]The prompts we use are listed in Appendix A.

in the previous quarter. This regression is conducted in the entire time window prior to quarter t so that the coefficients ($\beta_{k,t}$) are known at quarter t without look-ahead bias. We then compose the weighted score at the national level, and for industry sector j:

$$AI\ Weighted\ Score_t = \sum_{k=1}^{14} \beta_{k,t} AI\ Score_{k,t}, \tag{3}$$

$$AI\ Ind.\ Weighted\ Score_{j,t} = \sum_{k=1}^{14} \beta_{k,t} AI\ Ind.\ Score_{k,j,t}. \tag{4}$$

Note that $AI\ Score_{k,t}$ and $AI\ Ind.\ Score_{k,j,t}$ are the k-th national AI score and the industry-j AI score in quarter t, respectively. We then use these AI weighted scores to predict economic growth at both the national and industry levels.

Table 11 presents the results. Panel A indicates that the *AI Weighted Score* outperforms the *AI Economy Score* in predicting future GDP growth at the national level. This is evidenced by its stronger economic significance, better fit, and the persistence of its predictive power for up to the 14th quarter, whereas the *AI Economy Score* is predictive for only up to the 4th quarter (Table 3).

[Insert Table 11 Here]

Panel B of Table 11 shows that *AI Ind. Weighted Score* is predictive of industry-level GDP growth for up to 16 quarters, after controlling for other variables. Comparing column (1) of Panel B with column (1) of Table 6 Panel A reveals that R^2 in next-quarter industry GDP forecasts improves by 2.6% when including the *AI Ind. Weighted Score*, compared to 1.3% when *AI Ind. Score* is included. Moreover, the statistical significance and fit in long-term GDP growth prediction remain stronger up to the 16th quarter when using the *AI Ind. Weighted Score* relative to the *AI Ind. Score*.

Panel C reports the means and t-statistics of the time-series quarterly weights of the 14 scores in *AI Weighted Score*, i.e., $\beta_{k,t}$ as in (3). The results show that almost all of the weights are positive and statistically significant, except for those for customer demand and cost of capital.

The average weights for the significant scores range from 0.027 to 0.11. The most significant scores are those for revenue, production, wages, employment, industry prospects, and capital expenditure. This indicates that the multidimensional forecasts of managers during conference calls provide useful information for predicting firm sales, which in turn help with forecasting industry and national economic conditions.

Overall, the results indicate that incorporating a broader range of information from managers' earnings call discussions can significantly enhance the predictive power for GDP growth at different levels, especially for longer-term predictions.

7. Robustness Tests and Additional Analyses

This section conducts robustness tests of the previous results.

7.1. Horse Race with Survey Forecasts

We conduct a horse race between our *AI Economy Score* and the forecasts of real GDP growth by professional forecasters in the SPF. The SPF is a survey that provides economic forecasts and has been widely used by researchers as a metric for macroeconomic expectations (e.g., Coibion and Gorodnichenko, 2012, Coibion and Gorodnichenko, 2015, Bordalo, Gennaioli, and Shleifer, 2020). It contains quarterly forecasts of multiple economic indicators, including real GDP. For the survey conducted at quarter t, we take the median forecasts of real GDP by all the respondents for quarter $t+1$, then construct the forecasted GDP growth rate using the realized real GDP in quarter $t-1$ as the base, because it is the most recent available realized GDP number when economists make forecasts.

The results are presented in Table 12. The dependent variable is the log difference of real GDP for the next quarter, consistent with our baseline specification. In column (1), we include the *Term Spread*, *Real FFR*, and the SPF-Forecasted real GDP growth. In column (2), we replace the SPF forecast with the *AI Economy Score*. We include both measures simultaneously in column

(3). The R^2 comparison between columns (1) and (2) show that the in-sample fit for the model including the *AI Economy Score* is 32.5% higher than that for the model including the *SPF-Forecasted GDP*. When we include both measures in the model in column (3), the SPF forecast loses its predictive power, whereas the *AI Economy Score* remains significant with a similar estimated coefficient. This result indicates that the *AI Economy Score* is a stronger predictor of real GDP for the near future than survey-based forecasts.

[Insert Table 12 Here]

We also run a separate vector autoregression model, similar to the one in Section 4.3, to assess whether the AI Economy score provides additional predictive power beyond GDP forecasts by professionals. In Appendix Figure IA.3, we include GDP forecasts from the Survey of Professional Forecasters as an additional variable. Our results, shown in all subfigures of Figure 4, remain robust even when survey data is included. Notably, the SPF forecast significantly increases for the next four quarters following a positive shock to the AI Economy score, indicating a long-term impact on the professional forecasts from the AI Economy score.

7.2. Masked Identity Tests

Since the ChatGPT model is trained on a comprehensive dataset available up to September 2021, one potential concern is that in answering queries about conference calls, the model could have also used other information unavailable at the time of the calls. To address this potential "look-ahead bias," we conduct masked tests that remove all identifying information in the conference call transcripts, including firm information and dates. Glasserman and Lin (2023) demonstrate that look-ahead bias can be mitigated by removing company names from the prompt's text.

Specifically, we anonymize conference call texts by removing all dates, personal names, organization names, and product names. To do this, we first use regular expressions to identify years (from 1900 to 2099) and month names (including various abbreviations). Additionally, we

employ Python's spaCy library (en_core_web_sm model) to identify entity names of persons, organizations, and products. Once identified, each piece of identifying information is replaced with "###". To reduce the costs of performing the task, we limit the analysis to a random 10% subsample of our original final sample.

The results, presented in Table 13, show that *AI Economy Score_masked* still exhibits a positive and significant correlation with real GDP in the next quarter after controlling for known forecasters of GDP growth. The magnitudes of the coefficients are also comparable to those in Table 3. Results with predictions on other variables are presented in Table IA.4 of Appendix D. The *AI Economy Score_masked* demonstrates strong forecasting ability for industrial production, employment, and wages when constructed from the subsample with masked dates and other identifying information. These results largely alleviate concerns about look-ahead bias.

[Insert Table 13 Here]

7.3. Alternative Generative AI Model

In this section, we apply the same methodology as outlined in Section 3.1, but instead of using ChatGPT as the underlying language model, we utilize Meta Inc.'s open-source Llama-3 model. Llama-3 represents the latest advancements in open-source large language models and, in many respects, outperforms ChatGPT 3.5. The Llama-3 model is available in various sizes, ranging from 8 billion to 70 billion parameters. For our analysis, we specifically use the Meta-Llama-3-8B-Instruct model. Being open-source, Llama-3 allows researchers to run it on their own computers at no cost, making it more accessible.

Table 14 demonstrates that the results for Llama-3 are very similar to those obtained with ChatGPT. In Panel A, Column 2, the t-statistic is lower compared to Table 3, but remains significant at the 1% level. Likewise, the predictions remain significant for three quarters ahead in Panel B, as opposed to four quarters ahead in Table 3. The reduced statistical significance likely arises because ChatGPT, with its estimated 175 billion parameters, is somewhat more

advanced than the 8 billion-parameter model used for Llama-3. Nevertheless, the results in Table 14 suggest that our findings are not influenced by the choice of model.

8. Concluding Remarks

This paper explores the potentials of generative AI tools, specifically ChatGPT, to extract managerial expectations about the macroeconomy. We propose the *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 score has three major advantages: wide applicability, long-term predictability, and micro-level insights. First, the *AI Economy Score* can be readily constructed using publicly available data, offering an easily accessible set of expectation variables for researchers, policymakers, and investors. Second, the score exhibits forecasting power lasting up to ten quarters for a host of macroeconomic indicators, providing valuable insights for long-term economic decision-making. Furthermore, our methodology generates industry-level and firm-level forecasts, enabling analysis of economic expectations at a more granular level. These micro-level scores possess independent, strong predictive power that lasts even longer, for up to four years.

Our contribution extends beyond economic forecasting. By uncovering managerial expectations, we gain valuable insights into the thought processes and perspectives of top management regarding the future economic environment. This information can be instrumental for policymakers in crafting effective economic policies and for corporations in formulating strategic business decisions.

Future research directions could include exploring the integration of AI-generated forecasts with traditional economic models to enhance forecasting accuracy, investigating the impact of specific managerial characteristics and firm-level factors on AI Economy Scores, and examining

the applicability of our framework to other economic contexts and regions. As the field of AI continues to evolve, we anticipate further advancements in this area, leading to even more powerful tools for economic analysis and decision-making.

References

Barsky, R. B., and E. R. Sims. 2011. News shocks and business cycles. *Journal of Monetary Economics* 58:273–89. ISSN 0304-3932. doi:https://doi.org/10.1016/j.jmoneco.2011.03.001.

———. 2012. Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review* 102:1343–77.

Ben-David, I., J. R. Graham, and C. R. Harvey. 2013. Managerial miscalibration. *The Quarterly journal of economics* 128:1547–84.

Berger, P. G., and E. Ofek. 1995. Diversification's effect on firm value. *Journal of financial economics* 37:39–65.

Bernanke, B. S., and A. S. Blinder. 1992. The federal funds rate and the channels of monetary transmission. *The American Economic Review* 82:901–21. ISSN 00028282.

Bhandari, A., J. Borovička, and P. Ho. 2022. Survey data and subjective beliefs in business cycle models. *Available at SSRN 2763942* .

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni. 2013. News, noise, and fluctuations: An empirical exploration. *American Economic Review* 103:3045–70.

Bordalo, P., N. Gennaioli, and A. Shleifer. 2020. Memory, attention, and choice. *The Quarterly journal of economics* 135:1399–442.

Bybee, J. L. 2023. The ghost in the machine: Generating beliefs with large language models. *arXiv preprint arXiv:2305.02823* .

Candia, B., O. Coibion, and Y. Gorodnichenko. 2023. The macroeconomic expectations of firms. In *Handbook of Economic Expectations*, 321–53. Elsevier.

Chahrour, R., and K. Jurado. 2018. News or noise? the missing link. *American Economic Review* 108:1702–36.

Coibion, O., and Y. Gorodnichenko. 2012. What can survey forecasts tell us about information rigidities? *Journal of Political Economy* 120:116–59.

———. 2015. Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review* 105:2644–78.

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

Coibion, O., Y. Gorodnichenko, and T. Ropele. 2020. Inflation expectations and firm decisions: New causal evidence. *The Quarterly Journal of Economics* 135:165–219.

Coibion, O., Y. Gorodnichenko, and M. Weber. 2022. Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy* 130:1537–84.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding .

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *the Journal of Finance* 47:427–65.

Garcia, D., and Ø. Norli. 2012. Geographic dispersion and stock returns. *Journal of Financial Economics* 106:547–65.

Gernon, D. 2017. As trump disbanded advisory groups, this is who was in and who was out. https://www.cnbc.com/2017/08/16/heres-whos-in-and-out-of-trumps-economic-advisory-councils.html.

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

Glasserman, P., and C. Lin. 2023. Assessing look-ahead bias in stock return predictions generated by GPT sentiment analysis. Working Paper.

Graham, J. R., and C. R. Harvey. 2001. The theory and practice of corporate finance: Evidence from the field. *Journal of financial economics* 60:187–243.

Hansen, A. L., and S. Kazinnik. 2023. Can ChatGPT decipher Fedspeak? *Working paper, Available at SSRN 4399406* .

Jha, M., H. Liu, and A. Manela. 2022. Does finance benefit society? A language embedding approach. *Social Science Research Network. Working paper 3655263* .

Jha, M., J. Qian, M. Weber, and B. Yang. 2023. ChatGPT and corporate policies. *Working paper 32161, National Bureau of Economic Research* .

Kim, A., M. Muhn, and V. Nikolaev. 2024. Financial statement analysis with large language models. *Chicago Booth Research Paper* .

Kim, A. G., M. Muhn, and V. V. Nikolaev. 2023. Bloated disclosures: Can ChatGPT help investors process information? *Chicago Booth Research Paper* .

Korinek, A. 2023. Generative ai for economic research: Use cases and implications for economists. *Journal of Economic Literature* 61:1281–317.

Lang, L. H., and R. M. Stulz. 1994. Tobin's q, corporate diversification, and firm performance. *Journal of political economy* 102:1248–80.

Li, K., F. Mai, R. Shen, C. Yang, and T. Zhang. 2023. Dissecting corporate culture using generative AI – Insights from analyst reports. *Working paper, Available at SSRN 4558295* .

Lopez-Lira, A., and Y. Tang. 2023. Can ChatGPT forecast stock price movements? return predictability and large language models. *Working paper, University of Florida* .

Ouyang, S., H. Yun, and X. Zheng. 2024. How Ethical Should AI Be? How AI Alignment Shapes the Risk Preferences of LLMs. doi:10.2139/ssrn.4851711.

Saeedy, A. 2024. Jamie dimon warns u.s. might see interest-rate spike. https://www.wsj.com/finance/jamie-dimon-warns-u-s-might-face-interest-rate-spike-83789da7.

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

Schmitt-Grohé, S., and M. Uribe. 2012. What's news in business cycles. *Econometrica* 80:2733–64.

Sheng, J., Z. Sun, B. Yang, and A. L. Zhang. 2024. Generative AI and asset management. *Working Paper, Available at SSRN 4786575* .

van Binsbergen, J. H., S. Bryzgalova, M. Mukhopadhyay, and V. Sharma. 2024. (almost) 200 years of news-based economic sentiment. doi:10.3386/w32026.

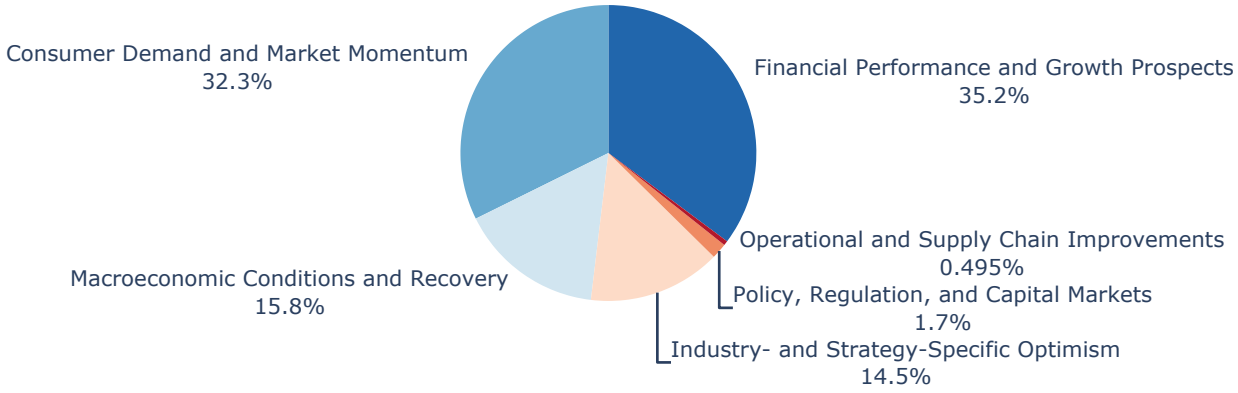
Weber, M., F. D'Acunto, Y. Gorodnichenko, and O. Coibion. 2022. The subjective inflation expectations of households and firms: Measurement, determinants, and implications. *Journal of Economic Perspectives* 36:157–84.

Yang, S. 2023. Predictive patentomics: Forecasting innovation success and valuation with ChatGPT. *Working paper, Available at SSRN 4482536* .

Figure 1. Explanations for Change in Optimism

We compile all the one-sentence explanations associated with positive (and separately, negative) changes in managerial optimism about the U.S. economy. We then prompt ChatGPT to group these explanations into distinct topics for each type of change.

(a) Explanations for Positive Change in Optimism



(b) Explanations for Negative Change in Optimism

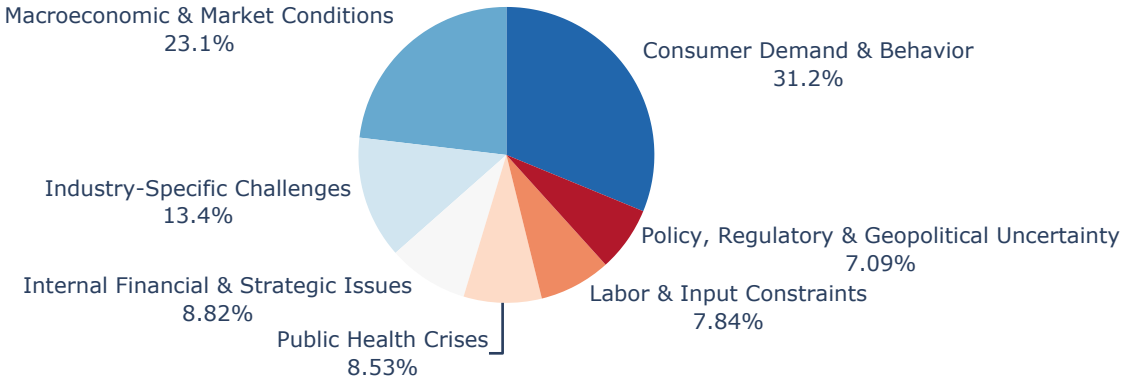
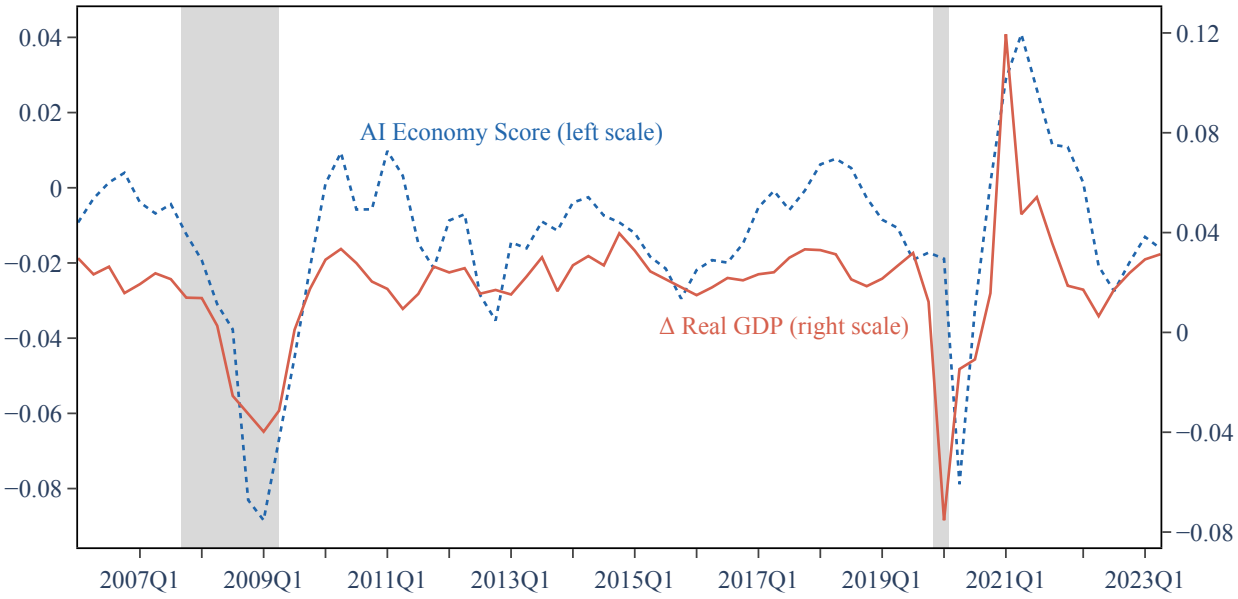


Figure 2. AI Economy Score vs. Realized and Forecasted GDP

The AI Economy Score is calculated by compiling responses to questions that ask, "Over the next quarter, how does the firm anticipate a change in optimism towards the US economy?" These responses are collected each quarter and aggregated to form a nationwide score. The changes in outcomes are (a) the percent change in realized GDP in quarter t+1 relative to quarter t-3, and (b) the change in Survey of Professional Forecasters (SPF)-based GDP forecast for quarter t+1 relative to the realized GDP in quarter t-3.

(a) *AI Economy Score* vs. Realized GDP Growth



(b) *AI Economy Score* vs. SPF Forecasts

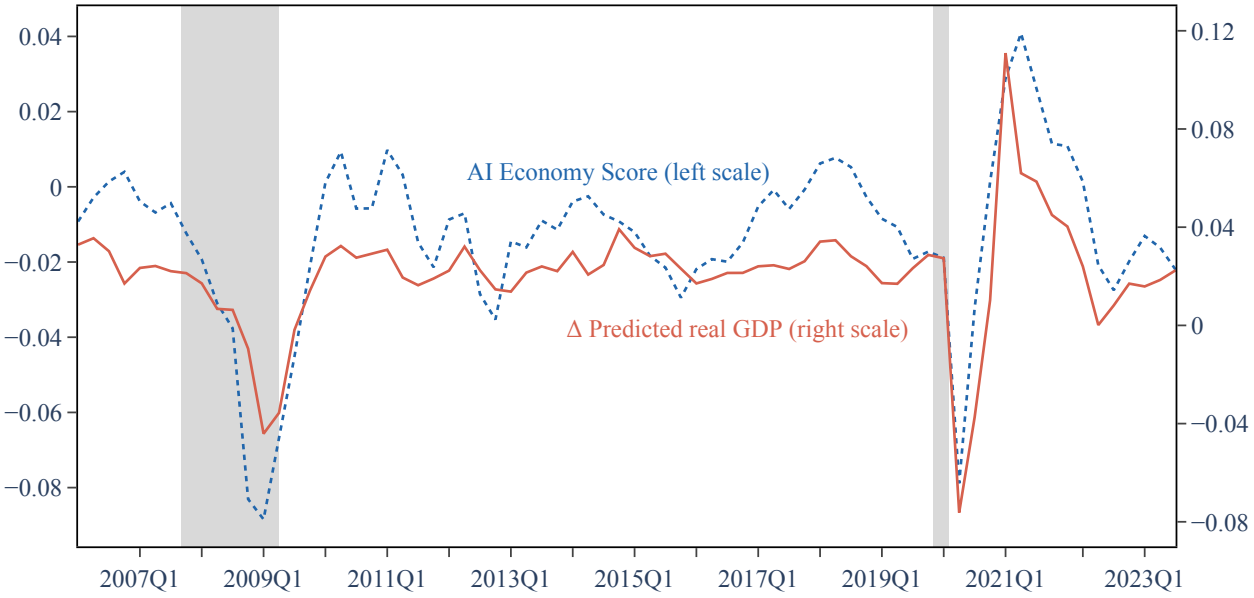


Figure 3. *AI Economy Score* across Industries

This figure represents average yearly AI Economy Score across industries. The score is calculated based on conference call texts of the firm (described in Section 3.1). The firms are aggregated into 19 industries, following the North American Industry Classification System (NAICS).

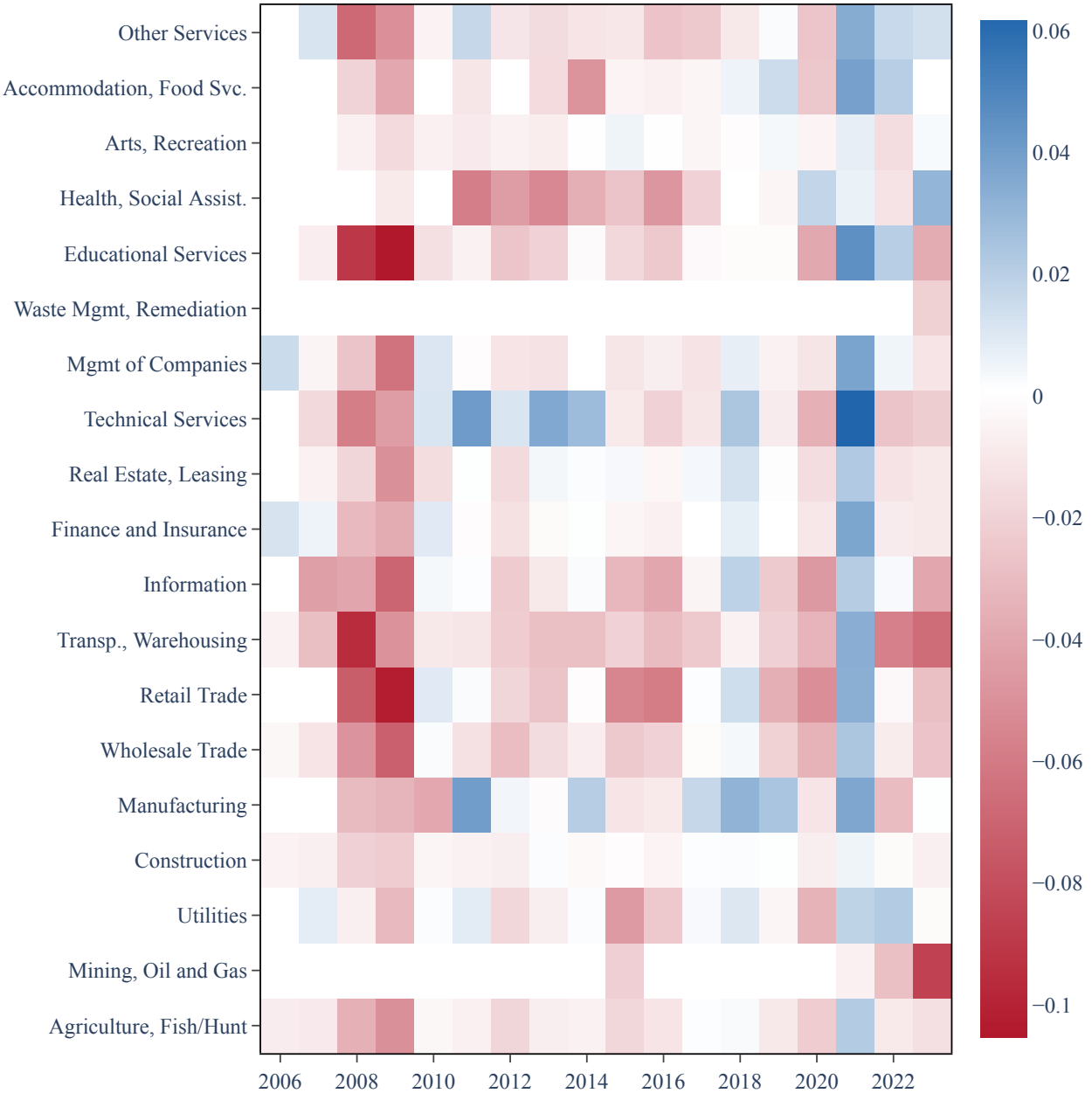


Figure 4. Macroeconomic implications of changes in *AI Economy Score*

The figure presents the impulse responses to a one-standard-deviation orthogonal shock to the *AI Economy Score*. The responses of consumption, investment, output growth, and excess market return have been accumulated. Shaded areas represent 95-percent confidence intervals, derived from 2,000 bootstrap replications. In each graph, the X-axis shows the quarters after the shock, and Y-axis is in percentage points.

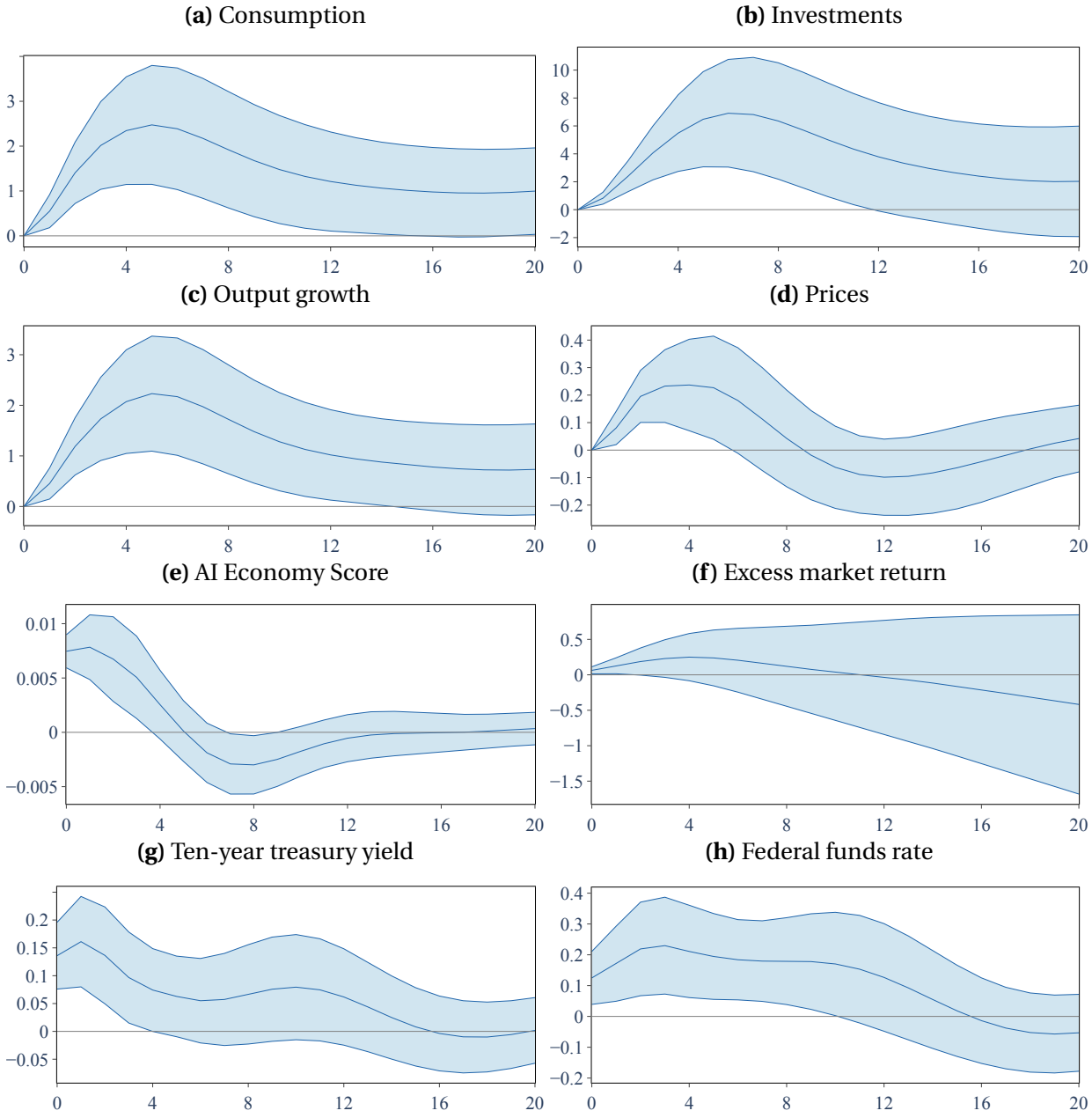


Figure 5. *AI Economy Score* across manager behavior types

This figure represents the time series trend of the AI Economy Score, aggregated from different groups of managers: Optimist, Pessimist, and Realist. A manager is classified as an Optimist (Pessimist) if the average bias of their past four-quarter forecasts ranks in the top (bottom) quartile, where bias is calculated as the tercile difference between their forecast and real GDP for each quarter. The remaining managers are classified as Realists.

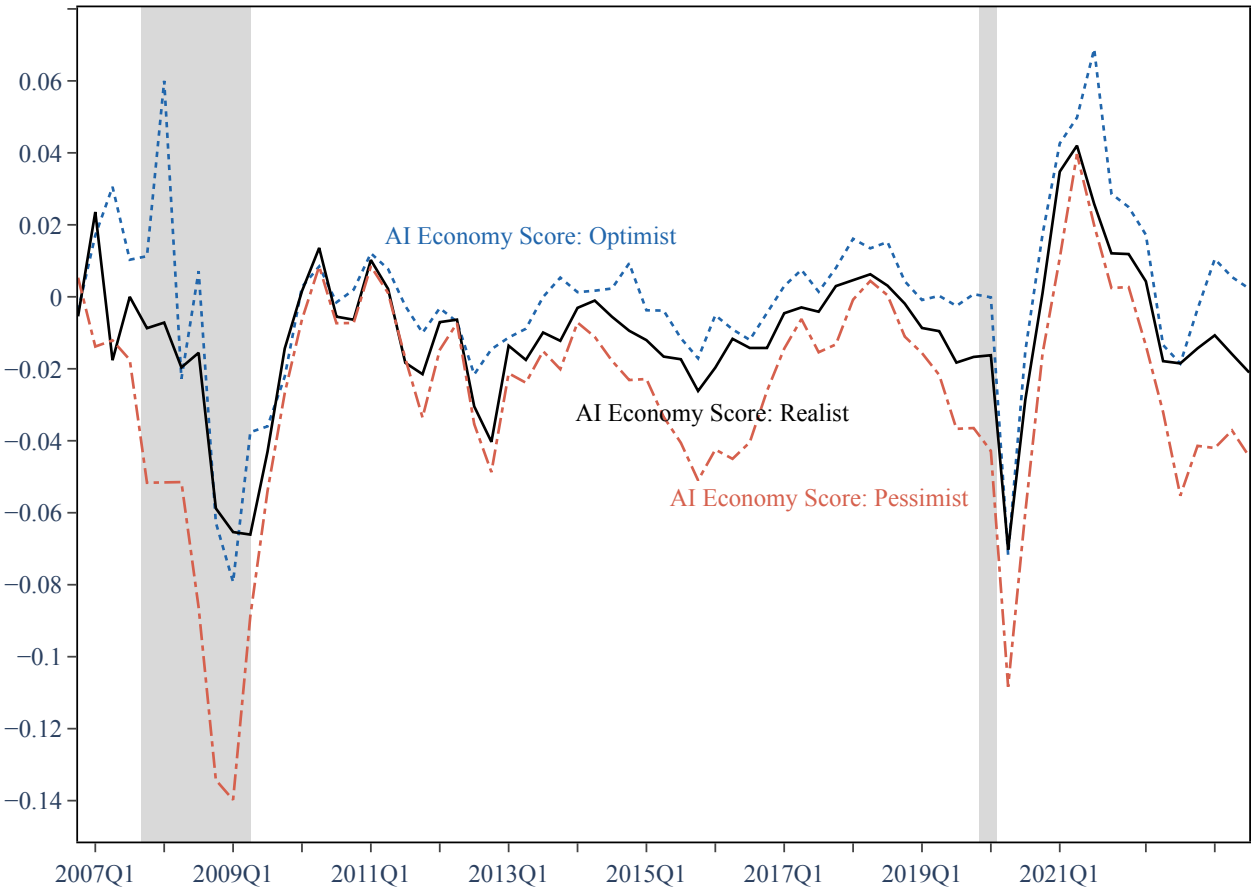


Table 1. Words associated with low- and high- *AI Economy Score*

This table displays the top ten 3-grams and 4-grams associated with both low and high AI Economy scores. We examine the reasons why ChatGPT assigned a significant decrease or decrease for low AI Economy scores (and similarly, a significant increase or increase for high AI Economy scores). To ensure consistency, we lemmatize each word to account for differing grammatical noun and verb forms. Additionally, we restrict the selection of words to those found in the English dictionary, excluding stop words and proper nouns.

Panel A: Top ten 3-grams

Low AI Economy Score	High AI Economy Score
challenge economic environment	firm express optimism
challenge market condition	strong revenue growth
lack thereof regard	strong financial performance
would likely lead	firm expect optimism
global economic condition	strong first quarter
current economic environment	firm remain optimistic
global economic environment	strong sale growth
natural gas price	company expect optimism
firm remain cautious	improve market condition
global economic growth	strong financial result

Panel B: Top ten 4-grams

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

Table 2. Summary Statistics: National Level

This table displays the descriptive statistics of AI-predicted managerial forecasts on economic indicators and the corresponding realized economic indicators at the national level. We obtain ChatGPT's responses on various aspects of economic activities based on earnings call transcripts, including forecasts on the US economy, quantity of products, number of employees, and wages. We convert the firm-level AI-predicted scores to aggregate-level scores by calculating the equally-weighted average across firms in each quarter. The national-level realized economic indicators are obtained from Federal Reserve Economic Data (FRED). We define the realized economic indicators as the logarithm of the next period (t+1) over the last period (t-1) in our regression analyses. *Term Spread* is the difference between ten-year and three-month constant-maturity Treasury yield. *Real FFR* is the real federal funds rate and the *GZ Spread* is the credit spread index. Variable definitions are in Appendix A.

	N	Mean	SD	p25	Median	p75
AI Economy Score	72	-0.013	0.022	-0.020	-0.010	-0.002
AI Economy Score_masked	72	.0004	0.017	-.006	.001	.008
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

Table 3. AI Economy Prediction and Real GDP Growth: National Level

This table presents the coefficients obtained by regressing the Real GDP for the next quarter/future quarters on AI-predicted scores at the national level. Panel A displays the results for the next quarter. The dependent variables are defined as the logarithm of the next period $t+1$ over the previous period $t-1$. Panel B reports the results for longer horizons. Here, the dependent variables are defined as the logarithm of the upcoming n-th quarter (up to six quarters) $t+n$ compared to the previous quarter $t-1$. Four lags of the dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

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

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

Table 4. Alternative Economic Indicators: National Level

This table presents the coefficients obtained by regressing the alternative economic predictors, *Industrial Production*, *Employment*, and *Wages*, for the next quarter on *AI Economy Score* at the national level. Panel A displays the results for predictions for the next quarter, while Panel B shows the results for long-term predictions. The dependent variables are defined as the logarithm of the upcoming n-th quarter (up to ten quarters) $t+n$ over the previous quarter $t-1$. Four lags of the dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

Panel A: Next Quarter

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Industrial Production			*Employment*			*Wages*		
	Forecating Period: Next Quarter			Forecating Period: Next Quarter			Forecating Period: Next Quarter		
Term Spread	1.964***	0.413	1.654**	0.588	0.0130	0.148	0.00237	-0.0853**	-0.108**
	(2.84)	(0.66)	(2.16)	(0.89)	(0.03)	(0.22)	(0.04)	(-2.37)	(-2.02)
Real FFR	0.601*	0.116	0.519	0.161	-0.0408	0.00573	-0.0192	-0.0533**	-0.0601**
	(1.70)	(0.33)	(1.42)	(0.46)	(-0.16)	(0.02)	(-0.54)	(-2.28)	(-2.22)
GZ Spread	-3.141***		-2.378***	-1.270***		-0.274	-0.172***		0.0399
	(-6.15)		(-4.02)	(-3.47)		(-0.42)	(-4.47)		(0.89)
AI Economy Score		1.334***	0.508**		0.641***	0.540*		0.109***	0.123***
		(4.93)	(2.44)		(5.19)	(1.86)		(8.01)	(5.27)
R-squared	0.520	0.410	0.547	0.288	0.361	0.365	0.790	0.855	0.857
Observations	72	72	72	72	72	72	72	72	72

(continued)

Panel B: Long-term

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Industrial Production				
	2 quarters	3 quarters	4 quarters	5 quarters	6 quarters	7 quarters	8 quarters	9 quarters	10 quarters
Term Spread	1.564**	1.467*	1.281	0.930	0.563	0.371	0.304	0.176	0.207
	(2.58)	(1.84)	(1.47)	(1.13)	(0.75)	(0.52)	(0.43)	(0.25)	(0.29)
Real FFR	0.296	-0.0188	-0.480	-1.077	-1.708**	-2.176***	-2.525***	-2.831***	-2.983***
	(0.84)	(-0.04)	(-0.86)	(-1.66)	(-2.38)	(-2.99)	(-3.75)	(-5.07)	(-6.92)
GZ Spread	-2.598***	-2.627***	-2.711***	-2.263**	-1.733**	-1.167	-0.975	-0.507	0.0323
	(-5.15)	(-3.90)	(-3.21)	(-2.54)	(-2.07)	(-1.52)	(-1.27)	(-0.64)	(0.04)
AI Economy Score	0.595**	0.680**	0.547	0.643*	0.699*	0.734*	0.558	0.571	0.759**
	(2.28)	(2.17)	(1.48)	(1.68)	(1.78)	(1.77)	(1.27)	(1.48)	(2.15)
R-squared	0.461	0.409	0.376	0.387	0.428	0.476	0.544	0.616	0.679
Observations	71	70	69	68	67	66	65	64	63

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Employment				
	2 quarters	3 quarters	4 quarters	5 quarters	6 quarters	7 quarters	8 quarters	9 quarters	10 quarters
Term Spread	-0.00000811	-0.0154	-0.0849	-0.130	-0.248	-0.410	-0.512	-0.726*	-0.861**
	(-0.00)	(-0.03)	(-0.14)	(-0.21)	(-0.43)	(-0.82)	(-1.17)	(-1.91)	(-2.52)
Real FFR	-0.161	-0.353	-0.623**	-0.929***	-1.300***	-1.670***	-1.959***	-2.233***	-2.449***
	(-0.60)	(-1.30)	(-2.27)	(-3.22)	(-4.24)	(-5.18)	(-5.95)	(-7.02)	(-8.32)
GZ Spread	-0.295	-0.429	-0.676	-0.931*	-1.066**	-0.947**	-0.994**	-0.935**	-0.758**
	(-0.64)	(-1.01)	(-1.51)	(-1.96)	(-2.05)	(-2.34)	(-2.47)	(-2.65)	(-2.05)
AI Economy Score	0.662***	0.757***	0.722***	0.642***	0.579**	0.603***	0.532**	0.465**	0.492**
	(3.00)	(3.71)	(3.67)	(3.17)	(2.57)	(3.03)	(2.35)	(2.19)	(2.49)
R-squared	0.377	0.418	0.430	0.456	0.500	0.549	0.599	0.634	0.664
Observations	72	71	70	69	68	67	66	65	64

(continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Wages				
	2 quarters	3 quarters	4 quarters	5 quarters	6 quarters	7 quarters	8 quarters	9 quarters	10 quarters
Term Spread	-0.201**	-0.342***	-0.544***	-0.747***	-1.002***	-1.208***	-1.462***	-1.661***	-1.852***
	(-2.56)	(-3.10)	(-3.78)	(-4.08)	(-4.74)	(-5.12)	(-5.65)	(-5.88)	(-5.96)
Real FFR	-0.123***	-0.211***	-0.322***	-0.446***	-0.600***	-0.750***	-0.905***	-1.074***	-1.251***
	(-2.83)	(-3.19)	(-3.78)	(-4.21)	(-4.78)	(-5.16)	(-5.69)	(-6.30)	(-6.85)
GZ Spread	0.0274	0.0350	0.0588	0.0157	0.0292	-0.0463	-0.0890	-0.144	-0.201
	(0.38)	(0.31)	(0.42)	(0.08)	(0.13)	(-0.18)	(-0.30)	(-0.49)	(-0.65)
AI Economy Score	0.171***	0.222***	0.269***	0.271**	0.298**	0.283*	0.272	0.273	0.276*
	(4.44)	(3.21)	(3.22)	(2.37)	(2.13)	(1.94)	(1.64)	(1.66)	(1.68)
R-squared	0.853	0.843	0.839	0.824	0.817	0.812	0.808	0.811	0.814
Observations	71	70	69	68	67	66	65	64	63

Table 5. Summary Statistics: Micro Level

This table displays the descriptive statistics of AI-predicted managerial forecasts on economic indicators and the corresponding realized economic indicators at the industry and firm levels, presented in Panels A and B, respectively. We obtain ChatGPT's responses on various aspects of economic activities based on earnings call transcripts, including forecasts on the US economy, quantity of products, number of employees, and wages. We convert the firm-level AI-predicted scores to industry-level scores by calculating the equally-weighted average within each of the 19 sectors under the NAICS classification standard (the sector *Other Services (except Public Administration)* is merged with *Public Administration (not covered in the economic census)*). The industry-level realized economic indicators are obtained from the Bureau of Economic Analysis (BEA), and the firm-level realized economic indicators are obtained from WRDS. We define the realized economic indicators as the logarithm of the next period (t+1) over the last period (t-1) in our regression analyses. Variable definitions are in Appendix A.

	N	Mean	SD	p25	Median	p75
Panel A: Industry-Level						
Quarterly						
AI Ind. Score	1,211	-0.012	0.035	-0.024	-0.005	0.004
Ind. Real GDP	1,211	0.009	0.054	-0.002	0.012	0.026
Annual						
AI Economy Score	279	-0.013	0.019	-0.020	-0.013	-0.003
AI Ind. Score	279	-0.012	0.027	-0.024	-0.007	0.001
Ind. Employment	279	0.011	0.073	-0.015	0.025	0.051
Ind. Wage	279	0.075	0.085	0.052	0.086	0.115
Panel B: Firm-Level Sample						
Quarterly						
AI Firm Score	77,305	0.341	0.214	0.250	0.400	0.500
Value-Add	77,305	0.029	0.406	-0.086	0.033	0.153
Sales	82,163	0.034	0.398	-0.061	0.033	0.135
Annually						
AI Firm Score	24,875	0.335	0.169	0.250	0.375	0.460
Firm Employment	24,875	0.065	0.380	-0.061	0.051	0.195
Firm Wage	2,102	0.122	0.269	0.000	0.118	0.246

Table 6. AI Economy Prediction and Real GDP Growth: Industry Level

This table reports the coefficients from regressing Real GDP for the next quarter/year on AI-predicted economy score at the industry level. We examine the predictive ability of the national level *AI Economy Score* and the score aggregated at the industry level, *AI Industry Score*. The *AI Industry Score* is obtained by taking the average of the firm-level scores within the 19 sectors based on the NAICS industry classification standard. Panel A presents the results for the next quarter; Panel B presents results for longer horizons; The dependent variables are defined as the logarithm of the next period $t+1$ over the last period $t-1$. Four lags of dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

Panel A: Next Period

	(1)	(2)	(3)	(4)
	Ind. Real GDP: Next Quarter			
Term Spread	0.958***	0.0728	-0.0805	0.596
	(2.86)	(0.26)	(-0.29)	(1.55)
Real FFR	0.142	-0.171	-0.222	0.0151
	(0.76)	(-1.02)	(-1.28)	(0.08)
GZ Spread	-1.792***			-1.007**
	(-7.78)			(-2.46)
AI Economy Score		0.783***		0.327
		(7.66)		(1.61)
AI Ind. Score			0.342***	0.120**
			(5.50)	(2.02)
Industry FE	Yes	Yes	Yes	Yes
R-squared	0.132	0.132	0.083	0.145
Observations	1,211	1,211	1,211	1,211

Panel B: Long Horizons

	(1)	(2)	(3)	(4)	(5)	(6)
	Ind. Real GDP: Long Horizons					
	2 quarters	3 quarters	4 quarters	8 quarters	12 quarters	16 quarters
Term Spread	0.412	0.612	0.514	-0.706*	-0.291	1.379***
	(1.21)	(1.50)	(1.24)	(-1.85)	(-0.73)	(2.79)
Real FFR	-0.326*	-0.629***	-1.155***	-3.010***	-2.792***	-1.475***
	(-1.76)	(-3.04)	(-5.13)	(-10.99)	(-10.36)	(-4.58)
GZ Spread	-0.960**	-1.405***	-1.641***	-0.969**	-0.465	-0.693
	(-2.50)	(-3.69)	(-4.28)	(-2.23)	(-1.02)	(-1.26)
AI Economy Score	0.477**	0.354*	0.223	0.0252	0.110	0.199
	(2.22)	(1.67)	(0.99)	(0.10)	(0.44)	(0.69)
AI Ind. Score	0.135**	0.181**	0.187**	0.252***	0.390***	0.330***
	(2.01)	(2.47)	(2.24)	(3.17)	(4.84)	(3.99)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.165	0.193	0.223	0.354	0.450	0.387
Observations	1,192	1,172	1,154	1,082	1,010	938

Table 7. Alternative Economic Indicators: Industry Level

This table reports the coefficients from regressing alternative economic predictors (excluding the real GDP) for the next year on AI-predicted scores at the industry level. We examine the predictive ability of the national level *AI Economy Score* and the score aggregated at the industry level, *AI Industry Score*. The *AI Industry Score* is obtained by taking the average of the firm-level scores within the 19 sectors based on the NAICS industry classification standard. Panel A presents the results on employment; Panel B presents the results on wages. The dependent variables are defined as the logarithm of the next period $t+1$ over the last period $t-1$. Two lags of dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

Panel A: Employment

	(1)	(2)	(3)	(4)
	Ind. Employment: Next Year			
Term Spread	3.289***	1.090	0.288	1.697**
	(4.11)	(1.58)	(0.42)	(2.14)
Real FFR	0.763*	-0.281	-0.449	-0.0143
	(1.88)	(-0.78)	(-1.21)	(-0.04)
GZ Spread	-4.815***			-1.249
	(-8.78)			(-1.20)
AI Economy Score		2.137***		1.144**
		(9.19)		(2.12)
AI Ind. Score			1.276***	0.571**
			(6.89)	(2.43)
Industry FE	Yes	Yes	Yes	Yes
R-squared	0.298	0.330	0.268	0.353
Observations	279	279	279	279

Panel B: Wages

	(1)	(2)	(3)	(4)
	Ind. Wages: Next Year			
Term Spread	1.358	-1.258	-2.196***	-0.302
	(1.40)	(-1.56)	(-2.76)	(-0.31)
Real FFR	0.128	-1.079**	-1.250***	-0.671
	(0.26)	(-2.44)	(-2.88)	(-1.43)
GZ Spread	-5.645***			-1.980*
	(-8.75)			(-1.81)
AI Economy Score		2.420***		0.920*
		(9.06)		(1.80)
AI Ind. Score			1.582***	0.866***
			(7.09)	(3.03)
Industry FE	Yes	Yes	Yes	Yes
R-squared	0.346	0.374	0.345	0.413
Observations	279	279	279	279

Table 8. AI Economy Prediction: Firm Level

This table reports the coefficients from regressing the economic predictors for the next period on AI-predicted scores at the firm level. In addition to national and industry-level AI-predicted scores on the US economy, we include the firm-level score in our analysis. The firm-level score, referred to as the [AI Firm Score], is derived from the manager's response to a question about the firm's prospects.Panel A presents the results on sales; Panel B presents results on earnings; Panel C presents results on employment; Panel D presents results on wages. The dependent variables are defined as the logarithm of the next period $t+1$ over the last period $t-1$. Control variables include *Size*, *Book-to-Market*, and *Tangibility*. Two lags of dependent variables are controlled for in all specifications Variable definitions are in Appendix A..

Panel A: Value-Add

	(1)	(2)	(3)	(4)	(5)
	Value-Add: Next quarter				
GZ Spread	-5.213***				-0.130
	(-6.15)				(-0.14)
AI Economy Score		2.483***			0.419
		(10.58)			(0.98)
AI Ind. Score			2.231***		1.414***
			(11.03)		(6.90)
AI Firm Score				0.310***	0.258***
				(10.75)	(15.37)
Controls	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.321	0.326	0.328	0.333	0.343
Observations	77,305	77,305	77,305	77,305	77,305

Panel B: Sales

	(1)	(2)	(3)	(4)	(5)
	Sales: Next quarter				
GZ Spread	-4.783***				-0.0212
	(-4.89)				(-0.02)
AI Economy Score		2.289***			0.716
		(9.94)			(1.55)
AI Ind. Score			1.981***		1.099***
			(10.85)		(5.90)
AI Firm Score				0.245***	0.194***
				(8.66)	(11.29)
Controls	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.272	0.277	0.279	0.279	0.288
Observations	82,163	82,163	82,163	82,163	82,163

(continued)

Panel C: Employment

	(1)	(2)	(3)	(4)	(5)
	Firm Employment: Next Year				
GZ Spread	-3.262***				-0.954
	(-5.81)				(-0.68)
AI Economy Score		1.212***			-0.304
		(4.53)			(-0.66)
AI Ind. Score			1.146***		0.465
			(5.06)		(1.22)
AI Firm Score				0.370***	0.352***
				(10.34)	(9.47)
Controls	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.357	0.357	0.358	0.373	0.374
Observations	24,875	24,875	24,875	24,875	24,875

Panel D: Wage

	(1)	(2)	(3)	(4)	(5)
	Firm Wages: Next Year				
GZ Spread	-5.074***				-0.569
	(-4.19)				(-0.42)
AI Economy Score		1.824***			-0.0398
		(5.05)			(-0.05)
AI Ind. Score			1.976***		1.332**
			(5.74)		(2.21)
AI Firm Score				0.316***	0.248***
				(6.63)	(4.88)
Controls	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.476	0.477	0.483	0.487	0.495
Observations	2,102	2,102	2,102	2,102	2,102

Table 9. AI-Predicted Economic Forecasts: Optimist vs. Pessimist

This table presents the AI economy prediction and realized GDP across different managerial behavior types. A manager is classified as an Optimist (Pessimist) if the average bias of their past four-quarter forecasts ranks in the top (bottom) quartile, where bias is calculated as the difference between their forecast and the real GDP for each quarter. The remaining managers are classified as Realists. Panel A shows the coefficients obtained by regressing the Real GDP for the next quarter on the AI-predicted score, aggregated by the groups of managers: Optimists, Pessimists, and Realists. The dependent variable is defined as the logarithm of the next period's GDP $(t+1)$ over the previous period's GDP $(t-1)$. Panel B reports the results at the firm level, where the *AI Economy Score* is a firm-specific score derived from a single earnings call. Optimist and Pessimist are indicators for managers classified as optimistic and pessimistic, respectively. The "Crisis" variable is an indicator of crisis periods in our sample, defined according to the NBER's criteria. It includes the last quarter of 2007, all four quarters of 2008, the first two quarters of 2009, and the first quarter of 2020. Four lags of the dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

Panel A: Macro Level

	(1)	(2)	(3)	(4)	(5)	(6)
	Real GDP: Next Quarter					
Term Spread	0.363	0.101	0.227	0.0669	0.241	0.0865
	(0.79)	(0.20)	(0.51)	(0.13)	(0.50)	(0.17)
Real FFR	0.0335	-0.0441	-0.0606	-0.0232	0.0137	-0.0606
	(0.14)	(-0.18)	(-0.26)	(-0.09)	(0.06)	(-0.27)
GZ Spread	-1.318***	-0.759*	-1.028***	-0.764*	-1.019***	-0.765*
	(-4.45)	(-1.91)	(-3.46)	(-1.95)	(-2.83)	(-1.98)
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

(continued)

Panel B: Firm Level

	(1)	(2)	(3)	(4)	(5)	(6)
			Real GDP: Next Quarter			
AI Economy Score	0.00430					
	(1.53)					
AI Economy Score#Optimist		0.00287	0.00196			0.00222
		(1.10)	(0.76)			(0.85)
AI Economy Score#Optimist#Crisis			0.0124			0.0151
			(0.37)			(0.44)
AI Economy Score#Pessimist				0.00587**	-0.000754	-0.000730
				(2.48)	(-0.18)	(-0.17)
AI Economy Score#Pessimist#Crisis					0.0852**	0.0854**
					(2.50)	(2.52)
Term Spread	0.806	0.808	0.807	0.807	0.776	0.775
	(1.02)	(1.03)	(1.02)	(1.02)	(1.00)	(0.99)
Real FFR	0.294	0.294	0.293	0.295	0.277	0.276
	(0.65)	(0.65)	(0.65)	(0.65)	(0.62)	(0.62)
GZ Spread	-2.011***	-2.020***	-2.018***	-2.015***	-1.959***	-1.956***
	(-2.70)	(-2.73)	(-2.71)	(-2.72)	(-2.69)	(-2.66)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.467	0.467	0.467	0.467	0.472	0.472
Observations	79,511	79,511	79,511	79,511	79,511	79,511

Table 10. AI-Predicted Economic Forecasts: Focused vs. Diversified Firms

This table presents the AI economy prediction and realized GDP across firms with different levels of operational diversification. *AI Economy Score* is a firm-specific score derived from a single earnings call. *Business Segment HHI* is the sum of the squared ratio of a business segment's sales to total sales. *High Business Segment HHI* is an indicator variable that represents whether the *Business Segment HHI* is above the median in a quarter across all firms. *GEO HHI* is the sum of the squared ratio of the number of times a state is mentioned in a 10-K filing to the number of different states mentioned. *High GEO HHI* is an indicator variable that represents whether the *GEO HHI* is above the median in a quarter across all firms. Four lags of the dependent variables are controlled for in all specifications. Standard errors are clusterred at the firm and year-quarter level. Variable definitions are in Appendix A.

	(1)	(2)	(3)	(4)
	Real GDP: Next Quarter			
AI Economy Score#Business Segment HHI	0.00609*			
	(1.85)			
AI Economy Score#High Business Segment HHI		0.00595**		
		(2.55)		
AI Economy Score#Geo HHI			0.0137**	
			(2.22)	
AI Economy Score#High Geo HHI				0.00684***
				(2.83)
Term Spread	0.836	0.838	0.836	0.837
	(1.06)	(1.06)	(1.06)	(1.06)
Real FFR	0.265	0.266	0.265	0.265
	(0.61)	(0.61)	(0.61)	(0.60)
GZ Spread	-1.894***	-1.898***	-1.895***	-1.897***
	(-2.89)	(-2.91)	(-2.90)	(-2.90)
Firm FE	Yes	Yes	Yes	Yes
R-squared	0.473	0.473	0.473	0.473
Observations	66,686	66,686	66,686	66,686

Table 11. AI Weighted Score and Real GDP Growth

This table presents the coefficients from regressions of real GDP for the next quarter or future quarters on weighted AI-predicted scores. The weighted AI-predicted score, AI Weighted Score, is derived from 14 scores corresponding to 14 questions posed to ChatGPT. The weight assigned to each score in each quarter is the beta loading estimated from past firm-quarter level observations, with firm sales as the dependent variable. Panels A and B display the results at the national and industry levels, respectively, while Panel C presents the mean and t-statistics of the time-series quarterly weights of the 14 scores. The t-statistics are based on Newey–West (1987) standard errors with two lags. All specifications control for four lags of the dependent variables. Variable definitions are in Appendix A.

Panel A: National Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
						Real GDP				
	1 qtr	2 qtrs	3 qtrs	4 qtrs	8 qtrs	12 qtrs	13 qtrs	14 qtrs	15 qtrs	16 qtrs
Term Spread	0.172	-0.108	-0.152	-0.287	-0.983***	-0.994***	-1.26***	-1.05***	-0.495	-0.399
	(0.40)	(-0.39)	(-0.43)	(-0.85)	(-6.51)	(-5.43)	(-5.26)	(-4.90)	(-0.87)	(-0.81)
Real FFR	-0.0480	-0.283*	-0.469***	-0.738***	-1.82***	-1.99***	-0.0210***	-0.0202***	-0.0176***	-0.0175***
	(-0.22)	(-1.86)	(-2.75)	(-4.01)	(-15.00)	(-14.38)	(-15.47)	(-14.01)	(-5.24)	(-5.98)
GZ Spread	-0.891***	-0.722***	-0.867***	-0.935***	-0.494**	-0.568***	-0.315	-0.432	-0.811***	-1.09***
	(-3.51)	(-2.97)	(-3.70)	(-3.12)	(-2.08)	(-3.05)	(-1.06)	(-1.57)	(-3.77)	(-3.97)
AI Weighted Score	0.185***	0.297**	0.306***	0.263***	0.238**	0.183***	0.282***	0.289***	0.142	0.113
	(3.51)	(2.45)	(3.02)	(2.71)	(2.61)	(2.94)	(3.75)	(3.32)	(1.21)	(1.04)
R-squared	0.485	0.520	0.580	0.587	0.775	0.795	0.800	0.791	0.727	0.741
Observations	70	69	68	67	63	59	58	57	56	55

(continued)

Panel B: Industry Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Ind. Real GDP			
	1 qtr	2 qtrs	3 qtrs	4 qtrs	8 qtrs	12 qtrs	16 qtrs
Term Spread	0.734**	0.529*	0.650*	0.477	-1.056***	-0.477	1.358***
	(2.13)	(1.69)	(1.73)	(1.27)	(-3.02)	(-1.32)	(2.92)
Real FFR	0.0154	-0.361**	-0.732***	-1.322***	-3.392***	-2.938***	-1.511***
	(0.08)	(-1.97)	(-3.55)	(-5.89)	(-13.13)	(-11.29)	(-4.81)
GZ Spread	-1.143***	-1.036***	-1.318***	-1.398***	-0.202	0.00298	-0.487
	(-3.84)	(-3.28)	(-4.30)	(-4.76)	(-0.59)	(0.01)	(-1.18)
AI Weighted Score	0.0372	0.145	0.106	0.0180	0.125	0.145	0.148
	(0.45)	(1.32)	(0.90)	(0.15)	(0.96)	(1.16)	(1.07)
AI Ind. Weighted Score	0.246***	0.276***	0.317***	0.384***	0.361***	0.402***	0.320***
	(5.24)	(4.67)	(4.45)	(5.14)	(4.48)	(5.31)	(4.09)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.158	0.179	0.208	0.244	0.377	0.456	0.393
Observations	1,203	1,184	1,164	1,146	1,074	1,002	930

Panel C: Summary Statistics of Weight of Scores

Scores	Mean of the Weight	T-stat
AI Earnings Score	0.0225***	(4.23)
AI Revenue Score	0.110***	(21.85)
AI Wage Score	0.0870***	(17.35)
AI Demand Score	0.00239	(0.95)
AI Economy Score	0.0366**	(2.05)
AI Global Score	0.0664***	(4.49)
AI Firm Score	0.0320***	(5.98)
AI Ind Score	0.0272***	(10.33)
AI Production Score	0.0640***	(17.21)
AI Product Price Score	0.0392***	(6.45)
AI Input Price Score	0.0453***	(6.38)
AI CostofCapital Score	-0.00493	(-0.33)
AI Capx Score	0.0384***	(7.02)
AI Employee Score	0.0572***	(10.69)

Table 12. AI Economic Forecast vs. Survey Forecast: A Horse Race

This table presents the coefficients from regressing next year's Real GDP on the AI-predicted economy score and the Survey-Forecasted economic growth at the national level. The survey-based measure, *SPF-Forecasted rGDP*, is derived from the Survey of Professional Forecasters. It uses the median forecasts of survey respondents for Real GDP in the upcoming quarter $(t+1)$ and the most recent realized RGDP for the previous quarter $(t-1)$ to construct a forecasted GDP growth. The dependent variables are defined as the logarithm of the next period $t+1$ over the last period $t-1$. Four lags of dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

	(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

Table 13. AI Economy Forecast: Masked Test

This table presents the coefficients from regressing next year's Real GDP on the AI-predicted economy score in a 10 percent subsample, with dates, firm names, person names, and product identifiers masked. The dependent variables are defined as the logarithm of the next period $t+1$ over the last period $t-1$. Four lags of dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

	(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

Table 14. Comparison with Alternative Generative AI Model

This table presents the coefficients obtained by regressing the Real GDP for the next quarter/future quarters on the AI-predicted score obtained from the Llama-3 model at the aggregate level. Panel A displays the results for the next quarter. The dependent variables are defined as the logarithm of the next period $t+1$ over the last period $t-1$. Panel B reports the results for longer horizons. Here, the dependent variables are defined as the logarithm of the upcoming n-th quarter (up to six quarters) $t+n$ compared to the previous quarter $t1$. Four lags of dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

Panel A: Next Quarter

	(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

Panel B: Long horizons

	(1)	(2)	(3)	(4)	(5)
			Real GDP		
	2 quarters	3 quarters	4 quarters	5 quarters	6 quarters
Term Spread	0.161	0.159	0.0202	-0.157	-0.433*
	(0.44)	(0.37)	(0.06)	(-0.51)	(-1.96)
Real FFR	-0.0174	-0.194	-0.467**	-0.741***	-1.056***
	(-0.09)	(-1.01)	(-2.47)	(-3.49)	(-4.59)
GZ Spread	-1.331***	-1.535***	-1.533***	-1.471***	-1.257***
	(-8.16)	(-6.56)	(-5.61)	(-5.91)	(-5.86)
AI Economy Score_Llama	0.722***	0.455**	0.275	0.288	0.289
	(3.23)	(2.10)	(1.13)	(1.10)	(1.51)
R-squared	0.496	0.530	0.536	0.573	0.604
Observations	71	70	69	68	67

Appendix A: Definitions of Variables

We ask ChatGPT to respond to the question: "*Over the next quarter, how does the firm anticipate a change in ... based on chunks from earnings call transcripts.*" We collect responses on the following 14 aspects, including optimism about the global economy, the firm's financial prospects, the industry's financial prospects, earnings, revenue, investments, wages and salaries, number of employees, demand for their products or services, production quantity, product or service prices, input or commodity prices, and the cost of capital. Based on the model's responses, we assign scores of -1, -0.5, 0, 0.5, or 1 for each of the following categories: Substantial Decrease, Decrease, No Change, Increase, and Substantial Increase, respectively. We then calculate the average of these scores across multiple chunks of one earnings call.

Variable	**Definition**
AI Firm Score	We collect ChatGPT's responses to the question related to US economy and then convert it to a score using the above approach.
AI Economy Score	The equally-weighted average of the firm-level score obtained from the question related to US economy for each quarter t.
AI Ind. Score	The equally-weighted average of the firm-level score obtained from the question related to US economy for each quarter t. There are 19 industries based on the NAICS industry classification system, after merging the sector *Other Services (except Public Administration)* with the sector *Public Administration (not covered in economic census)*.
AI Economy Score_masked	The *AI Economy Score* obtained in a 10% subsample with dates, firm names, person names, and product identifiers masked.
AI Ecnomy Score_Llama	The equally-weighted average of the firm-level score obtained from the question related to US economy using the Llama3-8b model (developed by Meta) for each quarter t.
AI Firm [Variable] Score	We collect ChatGPT's responses to the three questions (excluding the one related to US economy) using the above approach. The three questions-number of employees, quantity of products, and wage—are associated with *AI Firm Employment Score*, *AI Firm Production Score*, and *AI Firm Wage Score*, respectively.
AI Ind. [Variable] Score	We first obtain the firm-level scores: *AI Firm Employment Score*, *AI Firm Production Score*, and *AI Firm Wage Score*. We then calculate the equally-weighted average of these scores at the industry level for each quarter t to obtain the corresponding industry-level scores.
AI Ind. Weighted Score	We first compute firm-level scores for the 14 aspects. Next, we calculate the equally weighted average of these scores for each quarter t to obtain the corresponding industry-level scores. The *AI Ind. Weighted Score* is then derived from the industry-level scores, where the weight assigned to each score in each quarter is the beta loading estimated from past firm-quarter-level observations, using firm sales as the dependent variable.

(continued)

Variable	Definition
AI Weighted Score	We first compute firm-level scores for the 14 aspects. Next, we calculate the equally weighted average of these scores for each quarter t to obtain the corresponding national-level scores. The *AI Ind. Weighted Score* is then derived from the national-level scores, where the weight assigned to each score in each quarter is the beta loading estimated from past firm-quarter-level observations, using firm sales as the dependent variable.
AI [Variable] Score	We first obtain the firm-level scores: *AI Firm Employment Score*, *AI Firm Production Score*, and *AI Firm Wage Score*. We then calculate the equally-weighted average of these scores for each quarter t to obtain the corresponding national-level scores.
Book-to-Market	Follow the definition in Fama and French (1992).
Firm EBIT	The logarithm of firms' earnings before taxes and interests in the next period $(t+1)$ over the last period $(t-1)$.
Firm Employment	The logarithm of firms' number of employees in the next period $(t+1)$ over the last period $(t-1)$.
Firm Sales	The logarithm of firms' sales over total assets in the next period $(t+1)$ over the last period $(t-1)$.
GZ Spread	The equally-weighted GZ Spread defined by Gilchrist and Zakrajšek (2012) within each quarter t.
Industrial Production	The logarithm of Industrial Production Index in the next period $(t+1)$ over the last period $(t-1)$.
Ind. Employment	The logarithm of Industrial Nonfarm Payroll Employment in the next period $(t+1)$ over the last period $(t-1)$.
Ind. RGDP	The logarithm of Industrial Real GDP in the next period $(t+1)$ over the last period $(t-1)$.
Ind. Wage	The logarithm of Industrial Wages in the next period $(t+1)$ over the last period $(t-1)$.
Payroll Employment	The logarithm of Nonfarm Payroll Employment in the next period $(t+1)$ over the last period $(t-1)$.
Real FFR	The equally-weighted monthly real federal funds rate within each quarter t.
Real GDP	The logarithm of Real GDP in the next period $(t+1)$ over the last period $(t-1)$.
Tangibility	The Compustat item Property, Plant, and Equipment - Total (Net), scaled by total assets at the end of the quarter.
Term Spread	The equally-weighted monthly difference between the three-month constant-maturity Treasury yield and the ten-year constant-maturity yield within each quarter t.
Size	The natural logarithm of total book assets at the end of the quarter.

(continued)

Variable	**Definition**
SPF-Forecasted rGDP	It is the real GDP Growth rate forecasted by the survey respondents. For the survey conducted at quarter t, we take the median forecasts of real GDP among all the respondents for quarter $t+1$, then construct the forecasted real GDP growth rate using the median forecast for $t+1$ minus the realized real GDP in quarter $t-1$, divided by the realized real GDP in quarter $t-1$.
Unemployment	The logarithm of Unemployment Rate in the next period $(t+1)$ over the last period $(t-1)$.
Value-Add	Difference between sales and cost of goods sold.

Appendix B: Examples of high and low *AI Economy Score*

Category	Example Texts from Conference Call Transcripts
Significantly Increase (Score=1*)*	"Demand for our premium cabin has been remarkable across all entities, with premium paid load factor and RASM exceeding 2019 levels. And I'd like to underscore, we see a strong demand environment this summer, and we're highly confident that that will continue going forward. Looking ahead, we feel great about the industry and what's to come for American." "We are more optimistic about automotive than we were – we ever were. We expect the revenues to grow very nicely, but not in the short term. As you know, automotive takes at least a couple of years, depending on the application. The interest level on our products is very high, but probably the most important, the conversion of an opportunity into a design win is a lot higher than we've ever seen in any other market." "The large projects overall are really strong. I would even say if you look at any macro indicator, ABI, Dodge and so forth, the amount of large projects that are starting with some of the stimulus to go, whether it's new factories going up, whether chips, data centers being driven by artificial intelligence, and that's just in the United States, [indiscernible] again, we do have a global presence here, really drives some of the optimism that we had with the forecast for double-digit growth." "We are raising our 2024 outlook to reflect the strong momentum with which we exited 2023. The accelerating pace of investment in workflow automation and interest in Gen AI positions us well on our journey to becoming the defining enterprise software company of the 21st century." "U.S. growth of 12.7% was well ahead of our expectations with elective procedure volumes recovering and procedure cancellation rates returning to pre-pandemic levels. " "First, is the clearly growing industry-wide box office up in North America by some 29% quarter-to-quarter versus last year, growing to more than $.7 billion in the quarter. "

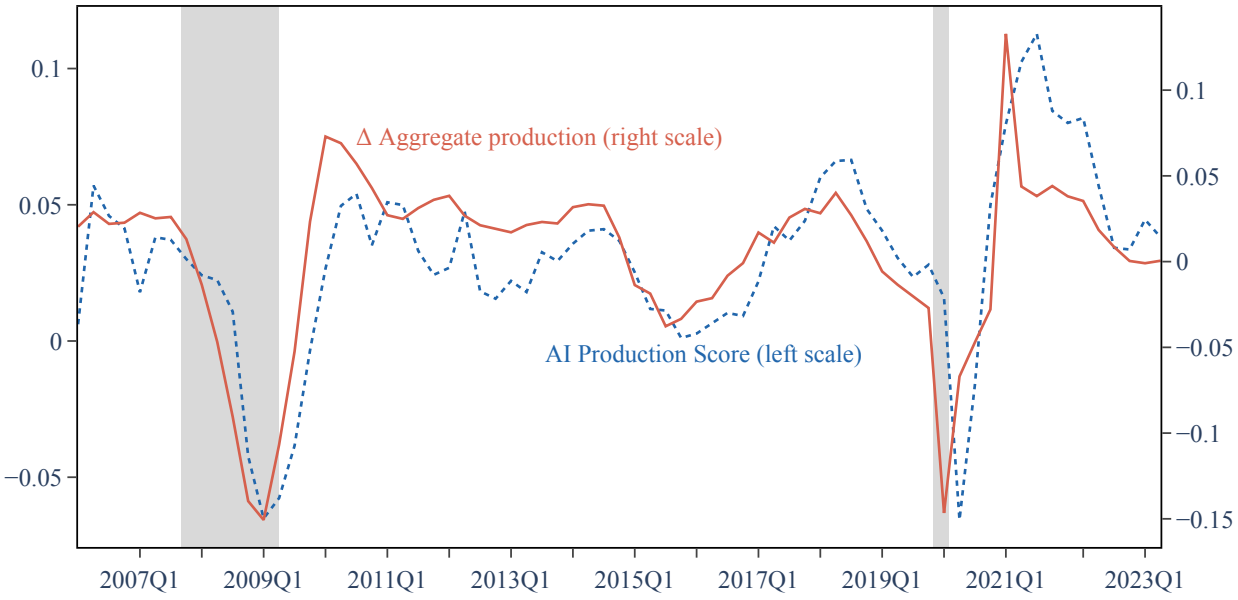
Category	Example Texts from Conference Call Transcripts
Significantly Decrease (Score=−1)	"As economic uncertainty persists, the strength of our model is enabling us to deliver value for our customers, continue to invest in our associates and deliver consistent shareholder return. Value remains top of mind for many of our customers as they are balancing several factors that are impacting their food at home spending. The effect of sustained inflation, reduced government benefits including SNAP, and higher interest rates have pressured customer spending, especially for those on a tight budget." "Recent results from across our industry and many of our largest customers has clearly demonstrated that macroeconomic headwinds continue largely unabated. This environment has challenged financial performance of IFF and our peers in the first half of 2023 and has resulted in a more cautious outlook for the remainder of the year." "Orders have softened in the residential home furnishings market. I think a lot of that is tied to macroeconomic uncertainty, Steven, I think, also with home sales slowing down. I think we'll continue to see a softening in demand." "Nonetheless, softness and natural gas pricing in the U.S. has had a dampening effect on current rig activity and is contributing to an increased level of contractual churn in the market, not only in terms of number of rigs but also the increased idle time between contracts." "Overall market declines in alternative protein consumption including demand decline for plant-based products, persistent supply chain challenges given the macro backdrop since the pandemic and more aggressive inventory management by customers have collectively contributed to pressure Functional Ingredients." "Across the globe, different regions are in varying phases of return to office compounded by varying economic conditions and a general slowdown in the housing market. Turning to Global Retail, as I mentioned earlier, we're seeing a slowdown in the housing market, particularly in luxury."

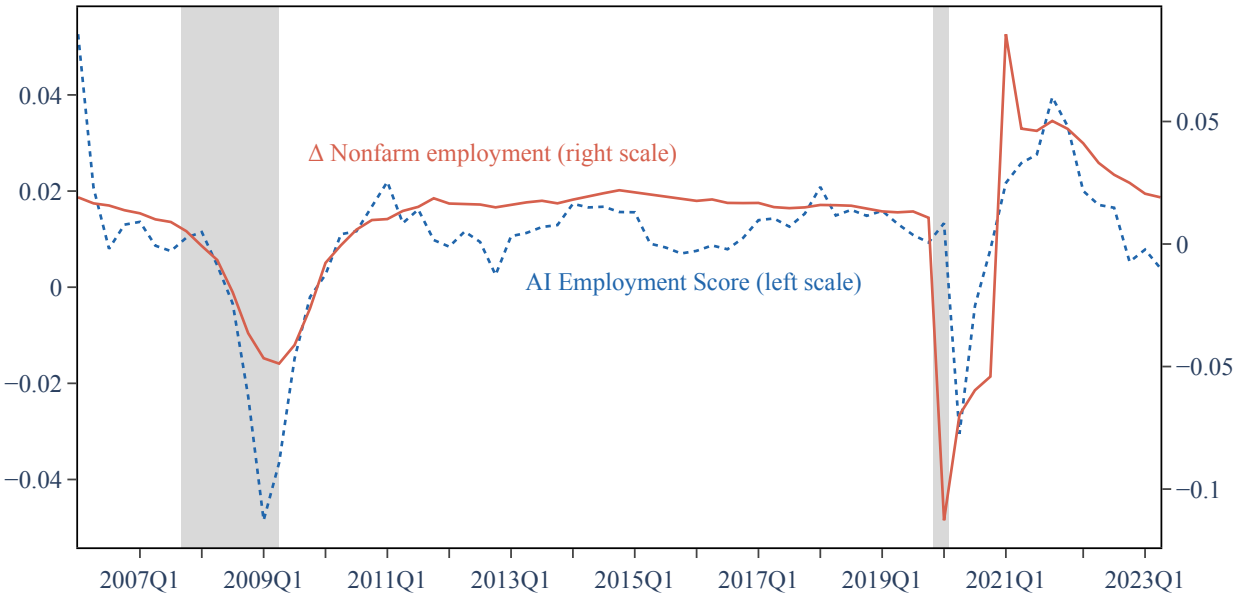
Internet Appendix of "Harnessing Generative AI for Economic Insights"

Figure IA.1. Additional Macroeconomic Variables and AI Scores

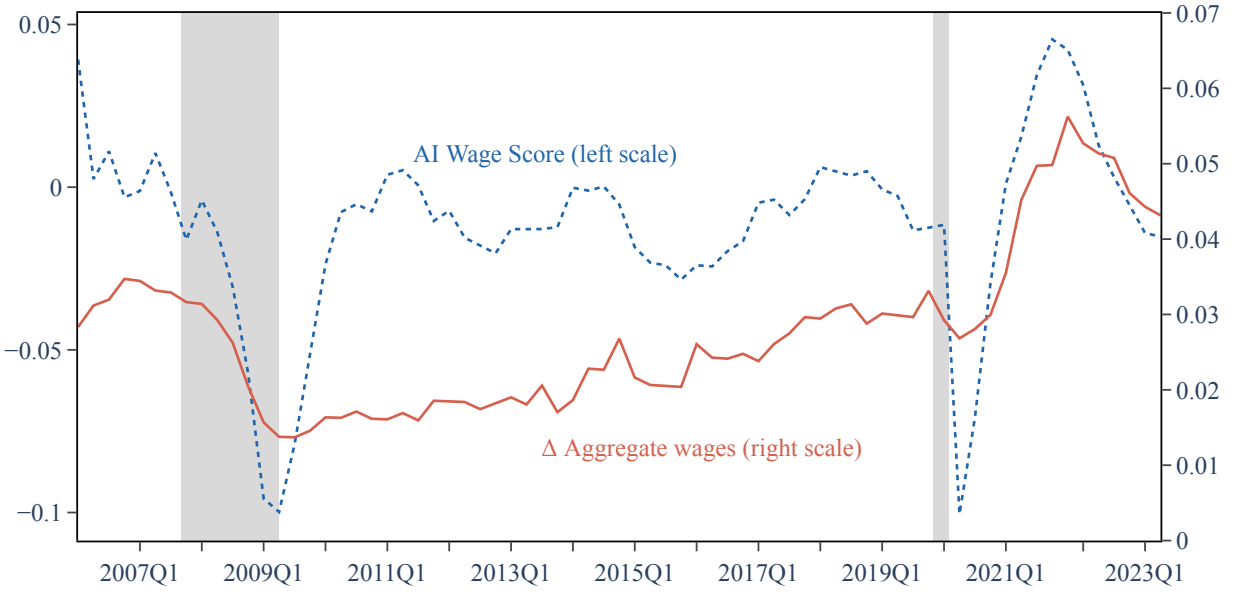
The ChatGPT score is calculated by compiling responses to questions that ask, "Over the next quarter, how does the firm anticipate a change in ..." These responses are collected each quarter and aggregated to form a nationwide ChatGPT score. The change in real outcomes are percent change in ... in quarter $t+1$ relative to quarter $t-3$. The change in predicted outcomes are percent change in ... in quarter $t+1$ relative to quarter $t-3$.

(a) Aggregate production and AI Production Score





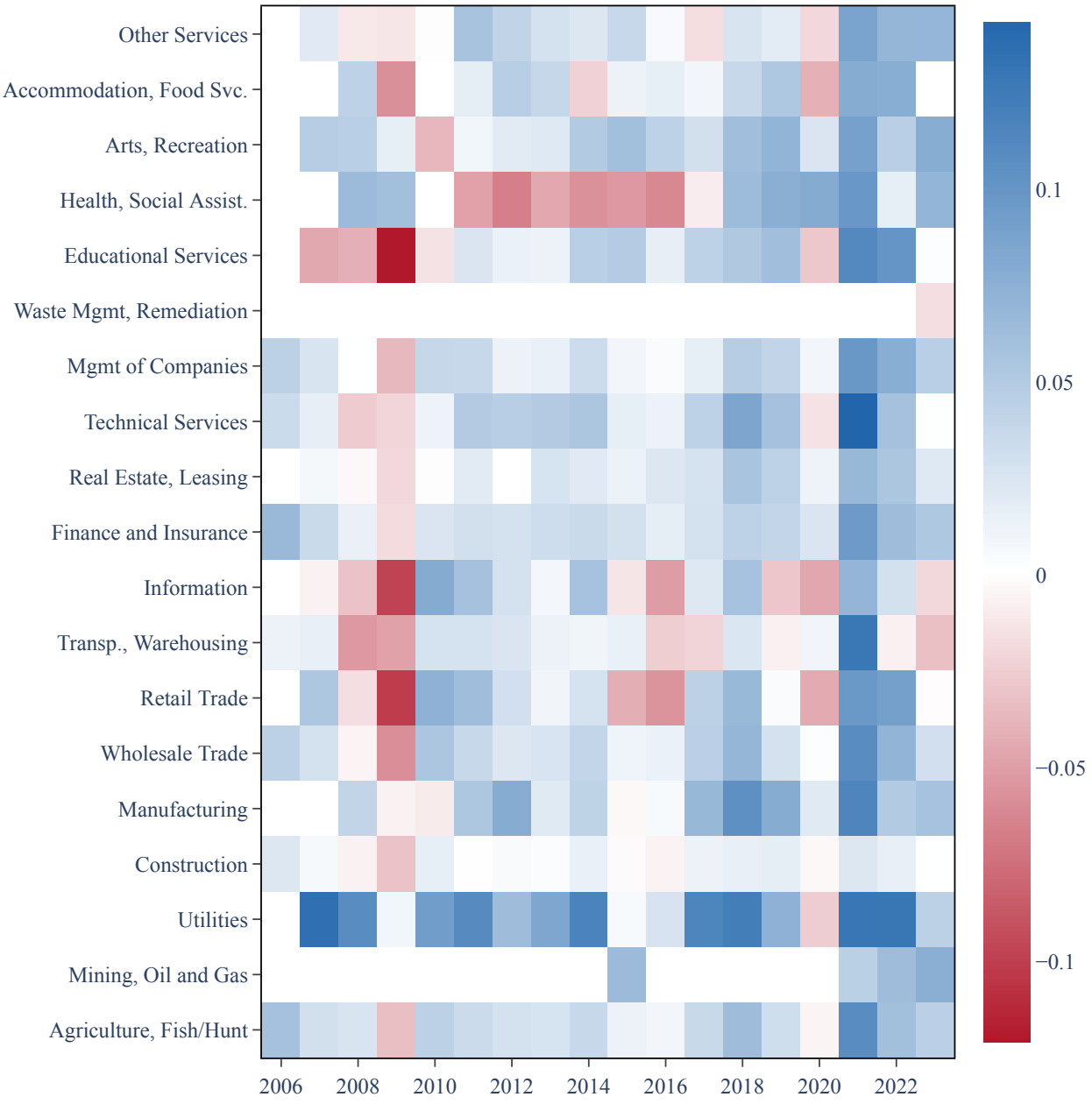
(b) Nonfarm payroll employment and AI Employment Score

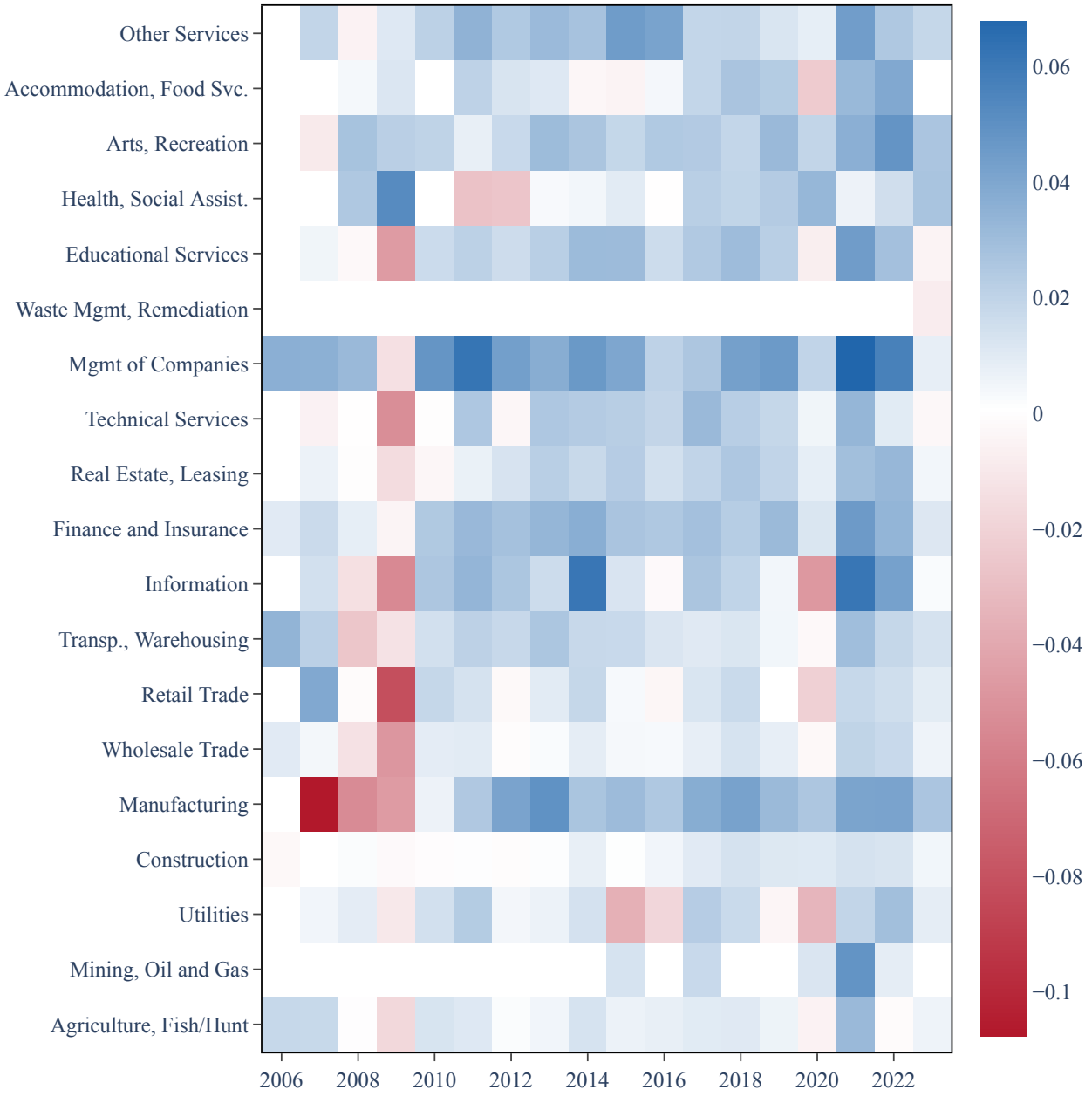


(c) Aggregate Wage and AI Wage Score

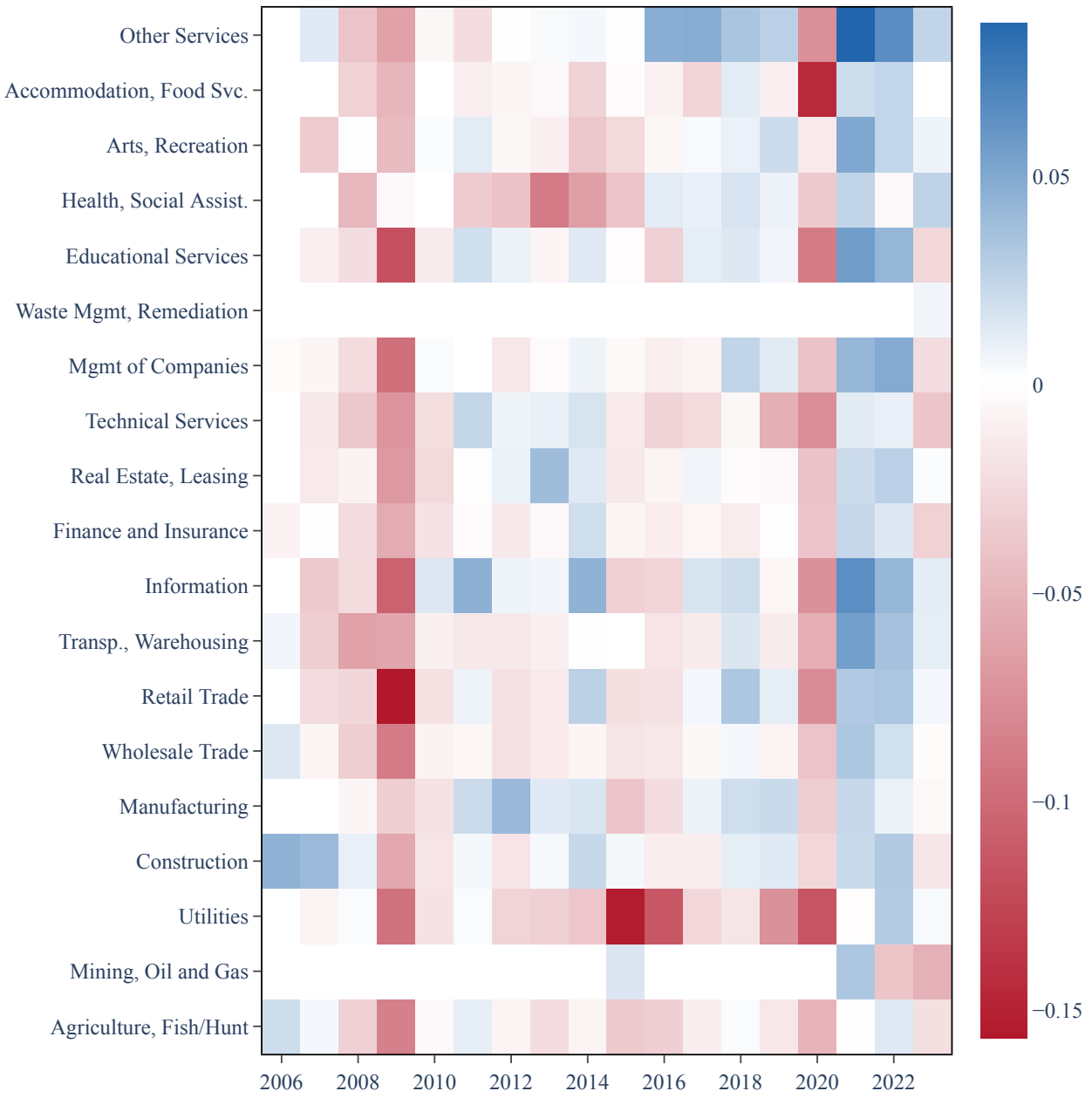
Figure IA.2. AI Production, Employment, and Wage Score across Industry Sectors

(a) AI Production Score





(b) AI Employment Score



(c) AI Wage Score

Figure IA.3. Macroeconomic implications of changes in *AI Economy Score*

The figure presents the impulse responses to a one-standard-deviation orthogonal shock to the *AI Economy Score*. The responses of consumption, investment, output growth, and excess market return have been accumulated. Shaded areas represent 95-percent confidence intervals, derived from 2,000 bootstrap replications. In each graph, the X-axis shows the quarters after the shock and Y-axis is in percentage points.

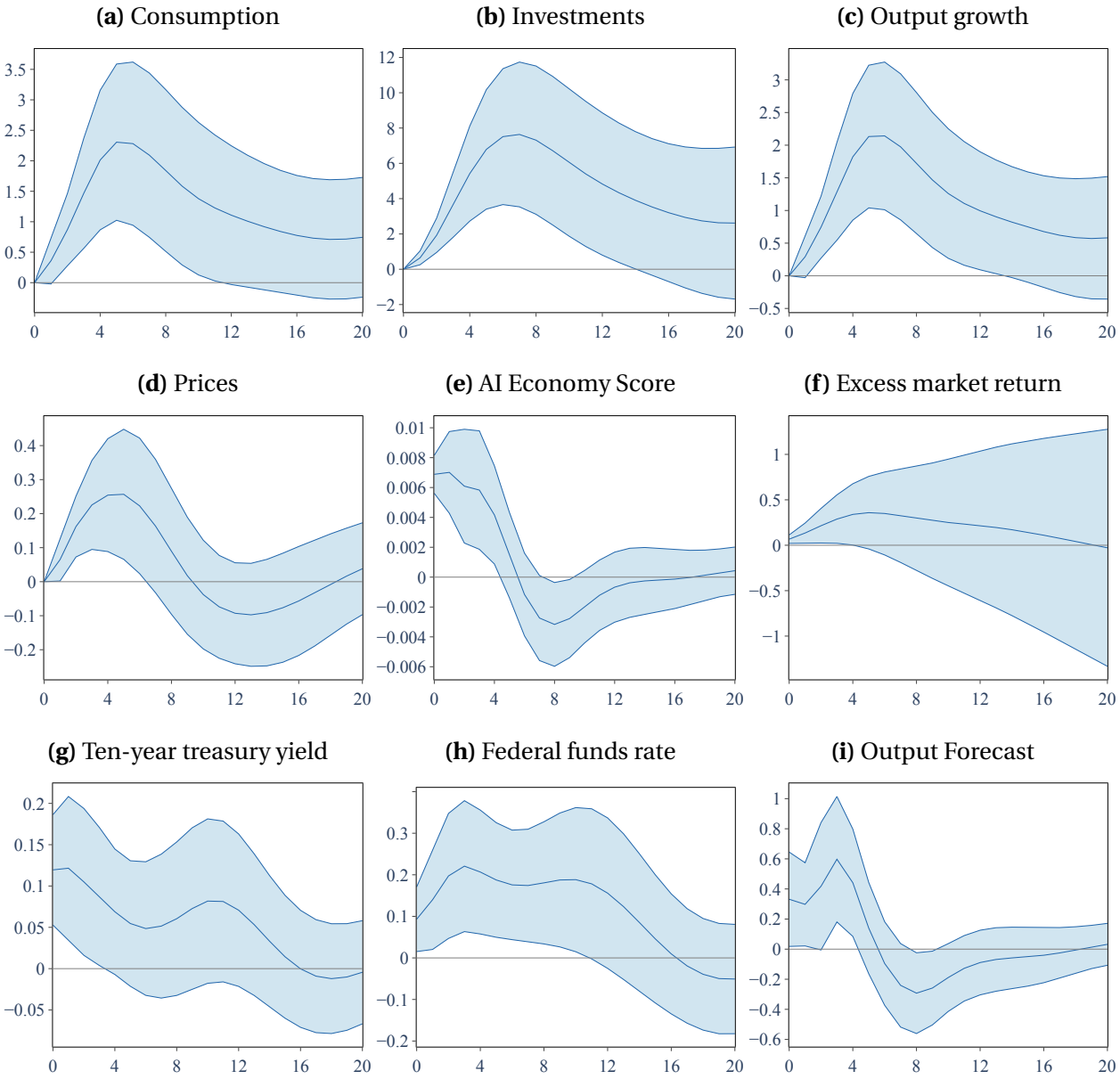


Table IA.1. AI predictions on other economic indicators: long horizons at the industry level

This table reports the coefficients from regressing the economic predictors for long horizons on AI-predicted scores at the industry level. Panel A displays the results on employment. Panel B displays the results on wages. The dependent variables are defined as the logarithm of the next period $t+1$ over the previous period $t-1$. Panel B reports the results for longer horizons. Four lags of the dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

Panel A: Industry Employment

	(1)	(2)	(3)
	Ind. Employment		
	2 years	3 years	4 years
Term Spread	2.109*	-0.219	3.992**
	(1.67)	(-0.16)	(2.24)
GZ Spread	-0.936	-2.662***	-0.814
	(-1.28)	(-3.39)	(-0.82)
Real FFR	-3.587*	-0.0647	-0.201
	(-1.80)	(-0.03)	(-0.09)
AI Economy Score	-0.645	0.285	1.029
	(-0.63)	(0.29)	(0.99)
AI Ind. Economy Score	0.717***	0.903***	1.012***
	(2.63)	(3.26)	(3.14)
Industry FE	Yes	Yes	Yes
R-squared	0.321	0.368	0.428
Observations	261	243	225
Industry FE	Yes	Yes	Yes
R-squared	0.321	0.368	0.428
Observations	261	243	225

(continued)

Panel B: Industry Wages

	(1)	(2)	(3)
		Ind. Wages	
	2 years	3 years	4 years
Term Spread	-2.087	-4.607***	-0.505
	(-1.45)	(-2.88)	(-0.24)
GZ Spread	-2.943***	-5.044***	-3.505***
	(-3.50)	(-5.62)	(-3.06)
Real FFR	-2.461	1.451	1.927
	(-1.14)	(0.60)	(0.73)
AI Economy Score	0.00287	1.452	2.746**
	(0.00)	(1.35)	(2.28)
AI Ind. Economy Score	1.078***	1.129***	1.373***
	(3.11)	(3.03)	(3.27)
Industry FE	Yes	Yes	Yes
R-squared	0.377	0.419	0.450
Observations	261	243	225

Table IA.2. AI Predictions for Long Horizons: Firm Level

This table reports the coefficients from regressing the economic predictors for long horizons on AI-predicted scores at the firm level. Panel A presents the results on value-add; Panel B presents results on firm sales; Panel C shows the results on firm employment; Panel D shows the results on firm wages. The dependent variables are defined as the logarithm of the next nth (up to four years) period $t+n$ over the last period $t-1$. Four (Two) lags of dependent variables are controlled for in quarterly (annually) specifications. Variable definitions are in Appendix A.

Panel A: Firm Value-Add

	(1)	(2)	(3)	(4)	(5)	(6)
				Value-Add		
	2 quarters	3 quarters	4 quarters	8 quarters	12 quarters	16 quarters
Term Spread	0.364	-1.613	-2.961**	-7.929***	-4.646***	-1.049
	(0.23)	(-1.04)	(-2.03)	(-5.41)	(-3.14)	(-0.53)
Real FFR	-1.149	-2.951**	-5.137***	-6.477***	-0.237	2.057
	(-1.07)	(-2.31)	(-4.06)	(-5.70)	(-0.22)	(1.26)
Book Leverage	0.0235	0.0329	0.0316	0.0527	0.0772	0.105
	(0.90)	(1.07)	(0.84)	(1.12)	(1.30)	(1.32)
Size	-0.0482***	-0.0813***	-0.113***	-0.233***	-0.340***	-0.412***
	(-6.02)	(-8.98)	(-9.89)	(-11.58)	(-12.97)	(-13.78)
Tangibility	-0.199***	-0.244***	-0.366***	-0.111	-0.307**	-0.416**
	(-2.71)	(-3.22)	(-4.24)	(-1.20)	(-2.50)	(-2.55)
GZ Spread	-1.486	-0.940	-0.823	1.216	-1.355	-0.398
	(-1.01)	(-0.83)	(-0.59)	(0.94)	(-0.82)	(-0.19)
AI Economy Score	-0.0345	-0.0740	-0.114	0.130	-1.558	0.343
	(-0.05)	(-0.11)	(-0.14)	(0.17)	(-1.56)	(0.30)
AI Ind. Score	1.319***	1.234***	1.130***	0.540*	1.022***	1.564***
	(7.44)	(6.45)	(4.59)	(1.93)	(4.41)	(5.82)
AI Firm Prospect Score	0.248***	0.233***	0.243***	0.193***	0.176***	0.188***
	(18.24)	(17.49)	(15.88)	(11.91)	(9.91)	(11.64)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.329	0.289	0.354	0.445	0.511	0.574
Observations	75,245	73,322	70,369	58,799	49,408	41,170

(continued)

Panel B: Firm Sales

	(1)	(2)	(3)	(4)	(5)	(6)
				Firm Sales		
	2 quarters	3 quarters	4 quarters	8 quarters	12 quarters	16 quarters
Term Spread	0.431	-1.545	-2.973*	-8.319***	-3.936**	-0.583
	(0.27)	(-0.99)	(-1.75)	(-4.77)	(-2.55)	(-0.28)
Real FFR	-0.568	-2.403**	-4.594***	-6.455***	0.716	3.276*
	(-0.57)	(-2.01)	(-3.39)	(-4.87)	(0.60)	(1.92)
Book Leverage	0.0276	0.0202	0.0145	0.0226	0.0310	0.0868
	(1.00)	(0.54)	(0.31)	(0.35)	(0.32)	(0.82)
Size	-0.0391***	-0.0683***	-0.0942***	-0.215***	-0.296***	-0.368***
	(-4.35)	(-7.38)	(-8.19)	(-10.60)	(-12.41)	(-13.26)
Tangibility	-0.271***	-0.336***	-0.396***	-0.154	-0.251**	-0.395***
	(-3.63)	(-4.32)	(-4.44)	(-1.59)	(-2.27)	(-2.90)
GZ Spread	-1.460	-0.131	-0.269	1.775	-0.627	0.221
	(-1.00)	(-0.10)	(-0.18)	(1.17)	(-0.37)	(0.10)
AI Economy Score	0.497	1.002	0.759	0.477	-1.553	0.407
	(0.70)	(1.41)	(0.87)	(0.53)	(-1.60)	(0.35)
AI Ind. Economy Score	1.061***	0.927***	0.840***	0.407	0.987***	1.707***
	(5.45)	(4.08)	(2.79)	(1.35)	(3.54)	(5.05)
AI Firm Prospect Score	0.207***	0.195***	0.220***	0.198***	0.185***	0.180***
	(15.04)	(14.40)	(14.95)	(10.80)	(11.08)	(10.49)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.292	0.274	0.328	0.433	0.506	0.603
Observations	80,671	78,788	75,747	63,355	53,056	44,326

(continued)

Panel C: Firm Employment

	(1)	(2)	(3)
	Firm Employment		
	2 years	3 years	4 years
Term Spread	-1.966	-3.310**	0.763
	(-1.47)	(-2.66)	(0.34)
Real FFR	-2.474***	-2.490***	0.645
	(-2.97)	(-3.46)	(0.43)
Book Leverage	-0.314***	-0.260***	-0.187**
	(-4.63)	(-4.02)	(-2.40)
Size	-0.106***	-0.202***	-0.269***
	(-4.24)	(-5.64)	(-6.25)
Tangibility	-0.557***	-0.639***	-0.708***
	(-5.21)	(-5.23)	(-4.89)
GZ Spread	-0.0818	0.765	-3.615**
	(-0.06)	(0.49)	(-2.17)
AI Economy Score	-0.663	-1.281	-2.988***
	(-1.01)	(-1.48)	(-3.49)
AI Ind. Economy Score	0.478	0.951**	1.603***
	(1.15)	(2.49)	(3.24)
AI Firm Prospect Score	0.341***	0.348***	0.316***
	(10.08)	(9.13)	(7.59)
Firm FE	Yes	Yes	Yes
R-squared	0.481	0.571	0.636
Observations	21,265	18,423	15,885

(continued)

Panel D: Firm Wages

	(1)	(2)	(3)
		Firm Wages	
	2 years	3 years	4 years
Term Spread	-3.389	-2.499	4.443
	(-1.13)	(-0.72)	(1.21)
Real FFR	-4.607*	-1.429	4.730
	(-1.79)	(-0.56)	(1.74)
Book Leverage	-0.171	-0.188	-0.135
	(-1.44)	(-1.48)	(-0.95)
Size	-0.0586	-0.124**	-0.185***
	(-1.52)	(-2.78)	(-3.73)
Tangibility	-0.578*	-0.743**	-0.747*
	(-1.97)	(-2.21)	(-1.85)
GZ Spread	1.889	-3.205	-9.286**
	(0.61)	(-0.68)	(-2.58)
AI Economy Score	0.874	-3.761	-5.090**
	(0.72)	(-1.40)	(-2.60)
AI Ind. Economy Score	0.219	1.075	1.816**
	(0.49)	(1.67)	(2.23)
AI Firm Prospect Score	0.248***	0.249***	0.269***
	(3.92)	(3.28)	(3.25)
Firm FE	Yes	Yes	Yes
R-squared	0.547	0.611	0.687
Observations	1,498	1,278	1,100

Table IA.3. AI Economy Prediction V.S. Survey Forecast: HorseRace on More Variables

This table presents the coefficients from regressing alternative economic predictors (excluding the real GDP) for the next quarter on the AI-predicted economy score and the Survey-Forecasted economic growth at the national level. The survey-based measure is obtained from the Survey of Professional Forecasters. It uses the median forecast of survey respondents for an economic indicator in the upcoming quarter (t+1) and the most recent actual value from the previous quarter (t-1) to construct a forecasted growth of the indicator. Panel A presents the results on industrial production; Panel B presents the results on employment. The dependent variables are the logarithm of the next period $t+1$ over the previous period $t-1$. Four lags of the dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

Panel A: Industrial Production

	(1)	(2)	(3)
	Industrial Production: Next Quarter		
Term Spread	0.545	-0.213	-0.170
	(0.78)	(-0.34)	(-0.23)
Real FFR	0.171	-0.0307	-0.0146
	(0.43)	(-0.08)	(-0.03)
AI Production Score	0.771***		0.0487
	(3.66)		(0.31)
SPF-Forecasted Production		1.332***	1.286***
		(3.60)	(3.42)
R-squared	0.256	0.440	0.440
Observations	72	72	72

Panel B: Employment

	(1)	(2)	(3)
	Employment: Next Quarter		
Term Spread	0.0542	0.0202	0.0537
	(0.11)	(0.05)	(0.13)
Real FFR	0.00135	0.00905	0.0187
	(0.00)	(0.04)	(0.08)
AI Employment Score	0.923***		0.269***
	(4.10)		(2.69)
SPF-Forecasted Employees		0.892***	0.740***
		(6.80)	(5.95)
R-squared	0.309	0.438	0.450
Observations	72	72	72

Table IA.4. AI Economy Prediction: Masked Test on More Variables

This table presents the coefficients from regressing next quarter's Industrial Production, Employment, or Wages on the AI-predicted economy score at the national level in a 10 percent subsample, with firm names, person names, and product identifiers masked. The dependent variables are defined as the logarithm of the next period $t+1$ over the previous period $t-1$. Four lags of the dependent variables are controlled for in all specifications. Variable definitions are in Appendix A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Industrial Production			*Employment*				*Wages*	
	Forecating Period: Next Quarter								
Term Spread	1.964***	0.719	1.783***	0.588	0.170	0.312	0.00237	-0.0719	-0.0522
	(2.84)	(1.10)	(2.80)	(0.89)	(0.33)	(0.58)	(0.04)	(-1.63)	(-0.91)
Real FFR	0.601*	0.459	0.670*	0.161	0.131	0.158	-0.0192	-0.0254	-0.0225
	(1.70)	(1.19)	(1.91)	(0.46)	(0.42)	(0.51)	(-0.54)	(-0.97)	(-0.81)
GZ Spread	-3.141***		-2.337***	-1.270***		-0.346	-0.172***		-0.0387
	(-6.15)		(-5.15)	(-3.47)		(-1.02)	(-4.47)		(-0.66)
AI Economy Score_masked		1.721***	0.740**		0.845***	0.685**		0.120***	0.103**
		(4.74)	(2.64)		(4.06)	(2.54)		(5.23)	(2.47)
R-squared	0.520	0.409	0.560	0.288	0.364	0.371	0.790	0.821	0.823
Observations	72	72	72	72	72	72	72	72	72